

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:02:07 ; Search time 102.067 Seconds  
(without alignments)  
3219.466 Million cell updates/sec

Title: US-09-830-972-2  
Perfect score: 5848  
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRKAD 1163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		8						
Result		Query						
No.	Score	Match	Length	DB	ID	Description		
1	5848	100.0	1163	5	ABB81074	Abb81074	Rat neuro	
2	5846	100.0	1163	3	AAY71310	Aay71310	Rat neuri	
3	5840	99.9	1162	3	AAY71557	Aay71557	Rat Nogo	
4	5823	99.6	1163	3	AAY71384	Aay71384	Alternati	
5	4921	84.1	974	3	AAY71560	Aay71560	Rat Nogo	
6	4403.5	75.3	1192	4	AAU04591	Aau04591	Human Nog	
7	4403.5	75.3	1192	5	ABP68600	Abp68600	Human pan	
8	4403.5	75.3	1192	6	ABR59667	Abr59667	Human Nog	
9	4398.5	75.2	1192	3	AAY56967	Aay56967	Human MAG	

10	4398.5	75.2	1192	4	AAB82349	Aab82349	Human	NOG
11	4398.5	75.2	1192	5	ABG30938	Abg30938	Human	Nog
12	4398.5	75.2	1192	5	ABB81078	Abb81078	Human	neu
13	4276.5	73.1	1178	3	AAAY71311	Aay71311	Human	neu
14	4116	70.4	1246	4	AAU33228	Aau33228	Novel	hum
15	4023	68.8	803	3	AAAY71562	Aay71562	Rat	Nogo
16	3714	63.5	737	3	AAAY71386	Aay71386	Rat	Nogo
17	3699.5	63.3	746	3	AAAY71391	Aay71391	Rat	Nogo
18	3651.5	62.4	736	3	AAAY71398	Aay71398	Rat	Nogo
19	3630.5	62.1	732	3	AAAY71399	Aay71399	Rat	Nogo
20	3494	59.7	695	3	AAAY71387	Aay71387	Rat	Nogo
21	3436	58.8	684	3	AAAY71394	Aay71394	Rat	Nogo
22	3385.5	57.9	983	6	ABU11573	Abu11573	Human	MDD
23	3280.5	56.1	893	3	AAAY95012	Aay95012	Human	sec
24	2779	47.5	552	3	AAAY71388	Aay71388	Rat	Nogo
25	2500.5	42.8	642	2	AAW58383	Aaw58383	Human	sec
26	2500.5	42.8	642	4	AAB90682	Aab90682	Human	BGL
27	2432	41.6	502	3	AAAY71396	Aay71396	Rat	Nogo
28	2388	40.8	475	3	AAAY71389	Aay71389	Rat	Nogo
29	2291	39.2	457	3	AAAY71392	Aay71392	Rat	Nogo
30	1987	34.0	403	3	AAAY71563	Aay71563	Rat	Nogo
31	1868	31.9	417	3	AAAY71393	Aay71393	Rat	Nogo
32	1801	30.8	356	3	AAAY71390	Aay71390	Rat	Nogo
33	1795.5	30.7	374	3	AAAY71397	Aay71397	Rat	Nogo
34	1513	25.9	379	7	ADB85283	Adb85283	Rat	fooce
35	1416	24.2	361	3	AAAY71385	Aay71385	Alternati	
36	1411.5	24.1	360	3	AAAY71383	Aay71383	Rat	neuri
37	1411.5	24.1	360	5	ABB81076	Abb81076	Rat	neuro
38	1405.5	24.0	359	3	AAAY71558	Aay71558	Rat	Nogo
39	1191	20.4	373	3	AAAY53624	Aay53624	A bone	ma
40	1191	20.4	373	3	AAAY56969	Aay56969	Human	MAG
41	1191	20.4	373	3	AAB24242	Aab24242	Human	Nog
42	1191	20.4	373	4	AAB82350	Aab82350	Human	NOG
43	1191	20.4	373	5	AAM47954	Aam47954	Human	RTN
44	1191	20.4	373	5	ABP68601	Abp68601	Human	pan
45	1191	20.4	373	5	ABB81079	Abb81079	Human	neu

#### ALIGNMENTS

##### RESULT 1

ABB81074

ID ABB81074 standard; protein; 1163 AA.

XX

AC ABB81074;

XX

DT 05-NOV-2002 (first entry)

XX

DE Rat neurotransmitter receptor protein Nogo-A.

XX

KW Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;  
KW central nervous system; peripheral nervous system; tranquillizer; Nogo;  
KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;  
KW nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;  
KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;  
KW neurotransmitter receptor; rat; receptor.

XX  
 OS Rattus norvegicus.  
 XX  
 PN US2002072493-A1.  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 28-JUN-2001; 2001US-00893348.  
 XX  
 PR 19-MAY-1998; 98IL-00124500.  
 PR 21-JUL-1998; 98WO-US014715.  
 PR 22-DEC-1998; 98US-00218277.  
 PR 19-MAY-1999; 99US-00314161.  
 XX  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 PI Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonego A;  
 PI Moalem G;  
 XX  
 DR WPI; 2002-607255/65.  
 DR N-PSDB; ABN86600.  
 XX  
 PT Promoting nerve regeneration and preventing neuronal degeneration in the  
 PT central/peripheral nervous system from injury/disease, comprises  
 PT administering nervous system-specific activated T cells/antigen, or  
 PT analogs/peptides.  
 XX  
 PS Example 5; Page 44-47; 93pp; English.  
 XX  
 CC The invention relates to promoting nerve regeneration or conferring  
 CC neuroprotection and preventing or inhibiting neuronal degeneration in the  
 CC central/peripheral nervous system (NS). The method involves administering  
 CC NS-specific activated T cells, NS-specific antigen, its analogue or its  
 CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or  
 CC combinations. The method is useful for promoting nerve regeneration and  
 CC preventing neuronal degeneration in central/peripheral nervous system  
 CC from injury/disease, where the injury is spinal cord injury, blunt  
 CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or  
 CC damages caused by surgery such as tumour excision. The disease is not an  
 CC autoimmune disease or neoplasm. The disease results in a degenerative  
 CC process occurring in either gray or white matter or both. The disease is  
 CC diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's  
 CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,  
 CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and  
 CC vitamin deficiency, intervertebral disc herniation, prion diseases such  
 CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral  
 CC neuropathies associated with various diseases, including but not limited  
 CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute  
 CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary  
 CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption  
 CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-  
 CC pathies, complications of various drugs (e.g., metronidazole) and toxins  
 CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia  
 CC telangiectasia, Friedreich's ataxia, amyloid polyneuropathies,  
 CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's  
 CC disease, or lipoproteinemia. The present sequence represents the rat  
 CC neurotransmitter receptor protein Nogo-A, an example of NS-specific

CC antigen  
XX  
SQ Sequence 1163 AA;

Query Match 100.0%; Score 5848; DB 5; Length 1163;  
Best Local Similarity 100.0%; Pred. No. 3.9e-297;  
Matches 1163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK	60
Db	1	MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK	60
Qy	61	PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	61	PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGLPAAPPAAPERQPSWERSPAAPAPSLP	120
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP	180
Db	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP	180
Qy	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Db	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Qy	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRCLEDSLEQKSLGKDSEGR	420
Db	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRCLEDSLEQKSLGKDSEGR	420
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIACDLIKETKLSTEPSDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	NAAVQETEAPYISIACDLIKETKLSTEPSDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780



Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Db	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Qy	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163
Db	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163

RESULT 2

AA71310

ID AAY71310 standard; protein; 1163 AA.

XX

AC AAY71310;

XX

DT 02-NOV-2000 (first entry)

XX

DE Rat neurite growth inhibitor Nogo A.

XX

KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS;

KW central nervous system; neoplastic disease; antiproliferative; glioma;

KW antisense gene therapy; neuroblastoma; meningioma; retinoblastoma;

KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;

KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;

KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;

KW structural plasticity; screening.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT Inhibitory-site 1. .171

FT /note= "Inhibits NIH 3T3 fibroblast spreading"

FT Modified-site 30

FT /note= "Casein kinase II site"

FT	Region	31. .58
FT		/note= "Acidic region"
FT	Region	31. .57
FT		/note= "Region specifically described in claim 16"
FT	Region	172. .259
FT		/note= "This region is not essential for inhibitory activity"
FT	Modified-site	233
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	242. .244
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	291
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	295
FT		/note= "Protein kinase C (PKC) site"
FT	Misc-difference	404
FT		/note= "Encoded by TTG"
FT	Modified-site	436
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	468. .470
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	484
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	488
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	502
FT		/note= "Casein kinase II site"
FT	Inhibitory-site	542. .722
FT	Modified-site	576
FT		/note= "Casein kinase II site"
FT	Peptide	623. .640
FT		/note= "used as immunogen to generate antibody AS 472"
FT	Modified-site	626
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	694. .696
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	715
FT		/note= "Casein kinase II site"
FT	Peptide	762. .1163
FT		/note= "used as immunogen to generate antibody AS Bruna"
FT	Modified-site	784
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	821
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	850
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	855
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	863
FT		/note= "Casein kinase II site"
FT	Modified-site	868
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	893
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	912. .914
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	925. .927

FT /note= "Asn is N-glycosylated"  
 FT Modified-site 954  
 FT /note= "PKC and casein kinase II sites"  
 FT Modified-site 956  
 FT /note= "PKC and casein kinase II sites"  
 FT Region 975. .1162  
 FT /note= "This region is not essential for inhibitory  
 FT activity"  
 FT Region 976. .1163  
 FT /note= "C-terminal common region found in Nogo A, B and C  
 FT isoforms"  
 FT Domain 988. .1023  
 FT /label= Transmembrane\_domain  
 FT /note= "C-terminal hydrophobic region specifically  
 FT described in claim 16"  
 FT Modified-site 1024  
 FT /note= "Protein kinase C (PKC) site"  
 FT Modified-site 1071. .1073  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1073  
 FT /note= "Protein kinase C (PKC) site"  
 FT Modified-site 1089  
 FT /note= "Protein kinase C (PKC) site"  
 FT Domain 1090. .1125  
 FT /label= Transmembrane\_domain  
 FT /note= "C-terminal hydrophobic region specifically  
 FT described in claim 16"  
 FT Modified-site 1141. .1143  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 1143  
 FT /note= "Protein kinase C (PKC) site"  
 XX  
 PN WO200031235-A2.  
 XX  
 PD 02-JUN-2000.  
 XX  
 PF 05-NOV-1999; 99WO-US026160.  
 XX  
 PR 06-NOV-1998; 98US-0107446P.  
 XX  
 PA (SCHW/) SCHWAB M E.  
 PA (CHEN/) CHEN M S.  
 XX  
 PI Schwab ME, Chen MS;  
 XX  
 DR WPI; 2000-400052/34.  
 DR N-PSDB; AAD01173.  
 XX  
 PT Nogo proteins and nucleic acids useful for treating neoplastic disorders  
 PT of the central nervous system and inducing regeneration of neurons.  
 XX  
 PS Claim 3; Fig 2A; 122pp; English.  
 XX  
 CC The present sequence is a rat Nogo A protein which is a potent neural  
 CC cell growth inhibitor and is free of all central nervous system (CNS)  
 CC myelin material with which it is natively associated. The protein was  
 CC derived from a cDNA generated by fusing R018U37-3, R1-3U21 cDNAs isolated



Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Db	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Qy	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163
Db	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163

RESULT 3

AAY71557

ID AAY71557 standard; protein; 1162 AA.

XX  
 AC AAY71557;  
 XX  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE Rat Nogo A truncated protein used in the construction of mutant Nogo-A.  
 XX  
 KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS;  
 KW central nervous system; neoplastic disease; antiproliferative; glioma;  
 KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
 KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
 KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
 KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
 KW structural plasticity; screening; mutant; mutein.  
 XX  
 OS Rattus sp.  
 XX  
 PN W0200031235-A2.  
 XX  
 PD 02-JUN-2000.  
 XX  
 PF 05-NOV-1999; 99WO-US026160.  
 XX  
 PR 06-NOV-1998; 98US-0107446P.  
 XX  
 PA (SCHW/) SCHWAB M E.  
 PA (CHEN/) CHEN M S.  
 XX  
 PI Schwab ME, Chen MS;  
 XX  
 DR WPI; 2000-400052/34.  
 XX  
 PT Nogo proteins and nucleic acids useful for treating neoplastic disorders  
 PT of the central nervous system and inducing regeneration of neurons.  
 XX  
 PS Example; Page; 122pp; English.  
 XX  
 CC The patent relates to neurite growth inhibitor Nogo which is free of all  
 CC central nervous system (CNS) myelin material with which it is natively  
 CC associated. Nogo proteins and fragments displaying neurite growth  
 CC inhibitory activity are used in the treatment of neoplastic disease of  
 CC the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
 CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
 CC oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and  
 CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
 CC Therapeutics which promote Nogo activity can be used to treat or prevent  
 CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
 CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
 CC used to inhibit production of Nogo protein to induce regeneration of  
 CC neurons or to promote structural plasticity of the CNS in disorders where  
 CC neurite growth, regeneration or maintenance are deficient or desired. The  
 CC animal models can be used in diagnostic and screening methods for  
 CC predisposition to disorders and to screen for or test molecules which can  
 CC treat or prevent disorders or diseases of the CNS. The present sequence  
 CC is a truncated form of rat Nogo A protein shown in AAY71310, which is  
 CC used in the construction of mutant Nogo-A. Nogo-A is composed of His-  
 CC tag/T7-tag/vector/Nogo-A sequence aa 1-1162. Nogo A deletion mutants were

CC used for mapping the inhibitory sites of Nogo protein. Major inhibitory  
CC region was identified in the Nogo A sequence from amino acids 172-974,  
CC particularly amino acids 542-722. In addition, N-terminal region 1-171  
CC was found to be inhibitory to NIH 3T3 fibroblast spreading. Note: The  
CC present sequence is not given in the specification but is derived from  
CC rat Nogo A sequence shown in AAY71310. SEQ ID numbers 35-42 are referred  
CC in claim 32 and SEQ ID NO: 29 in disclosure of the specification.  
CC However, the specification does not include sequences for these SEQ ID  
CC numbers  
XX  
SQ Sequence 1162 AA;

Query Match 99.9%; Score 5840; DB 3; Length 1162;  
Best Local Similarity 99.9%; Pred. No. 1e-296;  
Matches 1161; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Db	1	MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Qy	61	PAAGLSAAAVPPAAAAPLLDFFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	61	PAAGLSAAAVPPAAAAPLLDFFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP	180
Db	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP	180
Qy	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Qy	241	GNLSAVSSSEGTTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMSSFKGSPKGES	300
Db	241	GNLSAVSSSEGTTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMSSFKGSPKGES	300
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Qy	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCIEDSLEQKSLGKDSEGR	420
Db	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCIEDSLEQKSLGKDSEGR	420
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660

Db	601	 LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	 NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	 PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	 NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPFTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLLEPCDLSFKNIYPK	900
Db	841	 IIDEFPFTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLLEPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Db	901	 DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	961	 RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	 VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	1081	FLVDDLVDLSLKFAVLMWVFTTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Db	1081	 FLVDDLVDLSLKFAVLMWVFTTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Qy	1141	NKSVKDAMAKIQAKIPGLKRKA	1162
Db	1141	 NKSVKDAMAKIQAKIPGLKRKA	1162

RESULT 4

AAAY71384

ID AAY71384 standard; protein; 1163 AA.

XX

AC AAY71384;

XX

DT 02-NOV-2000 (first entry)

XX

DE Alternative version of rat neurite growth inhibitor Nogo A.

XX

KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS;  
 KW central nervous system; neoplastic disease; antiproliferative; glioma;  
 KW antisense gene therapy; neuroblastoma; meningioma; retinoblastoma;  
 KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
 KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
 KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
 KW structural plasticity; screening.



```

XX
OS   Rattus sp.
XX
FH   Key                               Location/Qualifiers
FT   Inhibitory-site 1. .171
FT                                   /note= "Inhibits NIH 3T3 fibroblast spreading"
FT   Modified-site   30
FT                                   /note= "Casein kinase II site"
FT   Region          31. .58
FT                                   /note= "Acidic region"
FT   Region          172. .259
FT                                   /note= "This region is not essential for inhibitory
FT                                   activity"
FT   Misc-difference 223
FT                                   /label= Unknown
FT                                   /note= "There is Leu at this position in the sequence
FT                                   shown in AAY71310"
FT   Modified-site   233
FT                                   /note= "Protein kinase C (PKC) site"
FT   Modified-site   242. .244
FT                                   /note= "Asn is N-glycosylated"
FT   Modified-site   291
FT                                   /note= "Protein kinase C (PKC) site"
FT   Modified-site   295
FT                                   /note= "Protein kinase C (PKC) site"
FT   Misc-difference 404
FT                                   /note= "There is Ile at this position in the sequence
FT                                   shown in AAY71310"
FT   Modified-site   436
FT                                   /note= "Protein kinase C (PKC) site"
FT   Modified-site   468. .470
FT                                   /note= "Asn is N-glycosylated"
FT   Misc-difference 469
FT                                   /label= Unknown
FT                                   /note= "There is Lys at this position in the sequence
FT                                   shown in AAY71310"
FT   Modified-site   484
FT                                   /note= "Protein kinase C (PKC) site"
FT   Modified-site   488
FT                                   /note= "Protein kinase C (PKC) site"
FT   Modified-site   502
FT                                   /note= "Casein kinase II site"
FT   Inhibitory-site 542. .722
FT   Modified-site   576
FT                                   /note= "Casein kinase II site"
FT   Peptide         623. .640
FT                                   /note= "used as immunogen to generate antibody AS 472"
FT   Modified-site   626
FT                                   /note= "Protein kinase C (PKC) site"
FT   Misc-difference 661
FT                                   /note= "There is Asn at this position in the sequence
FT                                   shown in AAY71310"
FT   Modified-site   694. .696
FT                                   /note= "Asn is N-glycosylated"
FT   Modified-site   715
FT                                   /note= "Casein kinase II site"
FT   Peptide         762. .1163

```

FT		/note= "used as immunogen to generate antibody AS Bruna"
FT	Modified-site	784
FT		/note= "Protein kinase C (PKC) site"
FT	Misc-difference	820
FT		/note= "There is Leu at this position in the sequence shown in AAY71310"
FT	Modified-site	821
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	850
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	855
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	863
FT		/note= "Casein kinase II site"
FT	Modified-site	868
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	893
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	912. .914
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	925. .927
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	954
FT		/note= "PKC and casein kinase II sites"
FT	Modified-site	956
FT		/note= "PKC and casein kinase II sites"
FT	Region	975. .1162
FT		/note= "This region is not essential for inhibitory activity"
FT	Region	976. .1163
FT		/note= "C-terminal common region found in Nogo A, B and C isoforms"
FT	Domain	988. .1023
FT		/label= Transmembrane_domain
FT		/note= "C-terminal hydrophobic region"
FT	Modified-site	1024
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	1071. .1073
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	1073
FT		/note= "Protein kinase C (PKC) site"
FT	Modified-site	1089
FT		/note= "Protein kinase C (PKC) site"
FT	Domain	1090. .1125
FT		/label= Transmembrane_domain
FT		/note= "C-terminal hydrophobic region"
FT	Modified-site	1141. .1143
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	1143
FT		/note= "Protein kinase C (PKC) site"
XX		
PN	WO200031235-A2.	
XX		
PD	02-JUN-2000.	
XX		
PF	05-NOV-1999;	99WO-US026160.
XX		

PR 06-NOV-1998; 98US-0107446P.

XX

PA (SCHW/) SCHWAB M E.

PA (CHEN/) CHEN M S.

XX

PI Schwab ME, Chen MS;

XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic disorders

PT of the central nervous system and inducing regeneration of neurons.

XX

PS Claim 3; Fig 13; 122pp; English.

XX

CC The present sequence is an alternative version of rat Nogo A protein  
CC which is a potent neural cell growth inhibitor and is free of all central  
CC nervous system (CNS) myelin material with which it is natively  
CC associated. Nogo proteins and fragments displaying neurite growth  
CC inhibitory activity are used in the treatment of neoplastic disease of  
CC the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
CC oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and  
CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
CC Therapeutics which promote Nogo activity can be used to treat or prevent  
CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
CC used to inhibit production of Nogo protein to induce regeneration of  
CC neurons or to promote structural plasticity of the CNS in disorders where  
CC neurite growth, regeneration or maintenance are deficient or desired. The  
CC animal models can be used in diagnostic and screening methods for  
CC predisposition to disorders and to screen for or test molecules which can  
CC treat or prevent disorders or diseases of the CNS. Note: The present  
CC sequence is an alternative version of the Nogo A sequence shown in Fig.  
CC 2A (see AAY71310). SEQ ID numbers 35-42 are referred in claim 32 and SEQ  
CC ID NO: 29 in disclosure of the specification. However the specification  
CC does not include sequences for these SEQ ID numbers

XX

SQ Sequence 1163 AA;

Query Match 99.6%; Score 5823; DB 3; Length 1163;

Best Local Similarity 99.7%; Pred. No. 7.9e-296;

Matches 1159; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60  
|  
Db 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60  
  
QY 61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120  
|  
Db 61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120  
  
QY 121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETFLFALP 180  
|  
Db 121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETFLFALP 180  
  
QY 181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL 240  
|

Db	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASXPSSLSTVSVFKEHGYL	240
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Db	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360
Qy	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSELEQSLGKDSEGR	420
Db	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSELEQSLGKDSEGR	420
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENXTDEKKIEERKA	480
Qy	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	MAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Db	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080

Qy 1081 FLVDDLVDSLKFAVLMWVFTYVGALFENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA 1140  
 |||  
 Db 1081 FLVDDLVDSLKFAVLMWVFTYVGALFENGLTLLILALISLFSIPVIYERHQVQIDHYLGIA 1140  
 Qy 1141 NKSVKDAMAKIQAKIPGLKRKAD 1163  
 |||  
 Db 1141 NKSVKDAMAKIQAKIPGLKRKAD 1163

RESULT 5

AA71560

ID AA71560 standard; protein; 974 AA.

XX

AC AA71560;

XX

DT 02-NOV-2000 (first entry)

XX

DE Rat Nogo A protein fragment used in the construction of mutant NiAext.

XX

KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS;  
 KW central nervous system; neoplastic disease; antiproliferative; glioma;  
 KW antisense gene therapy; neuroblastoma; menangioma; retinoblastoma;  
 KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
 KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
 KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
 KW structural plasticity; screening; mutant; mutein.

XX

OS Rattus sp.

XX

PN WO200031235-A2.

XX

PD 02-JUN-2000.

XX

PF 05-NOV-1999; 99WO-US026160.

XX

PR 06-NOV-1998; 98US-0107446P.

XX

PA (SCHW/) SCHWAB M E.

PA (CHEN/) CHEN M S.

XX

PI Schwab ME, Chen MS;

XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic disorders  
 PT of the central nervous system and inducing regeneration of neurons.

XX

PS Example; Page; 122pp; English.

XX

CC The patent relates to neurite growth inhibitor Nogo which is free of all  
 CC central nervous system (CNS) myelin material with which it is natively  
 CC associated. Nogo proteins and fragments displaying neurite growth  
 CC inhibitory activity are used in the treatment of neoplastic disease of  
 CC the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
 CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
 CC oligodendroglioma, menangioma, neuroblastoma or retinoblastoma and

CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
CC Therapeutics which promote Nogo activity can be used to treat or prevent  
CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
CC used to inhibit production of Nogo protein to induce regeneration of  
CC neurons or to promote structural plasticity of the CNS in disorders where  
CC neurite growth, regeneration or maintenance are deficient or desired. The  
CC animal models can be used in diagnostic and screening methods for  
CC predisposition to disorders and to screen for or test molecules which can  
CC treat or prevent disorders or diseases of the CNS. The present sequence  
CC is a fragment of rat Nogo A protein shown in AAY71310, which is used in  
CC the construction of mutant NiAext. The mutant is composed of His-tag/T7-  
CC tag/vector/Nogo-A sequence aa 1-974/T7-tag. Nogo A deletion mutants were  
CC used for mapping the inhibitory sites of Nogo protein. Major inhibitory  
CC region was identified in the Nogo A sequence from amino acids 172-974,  
CC particularly amino acids 542-722. In addition, N-terminal region 1-171  
CC was found to be inhibitory to NIH 3T3 fibroblast spreading. Note: The  
CC present sequence is not given in the specification but is derived from  
CC rat Nogo A sequence shown in AAY71310. SEQ ID numbers 35-42 are referred  
CC in claim 32 and SEQ ID NO: 29 in disclosure of the specification.  
CC However, the specification does not include sequences for these SEQ ID  
CC numbers

XX

SQ Sequence 974 AA;

Query Match 84.1%; Score 4921; DB 3; Length 974;  
Best Local Similarity 99.9%; Pred. No. 8.8e-249;  
Matches 973; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEDIDQSSLVSSSTDSPRPFPFAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Db	1	MEDIDQSSLVSSSTDSPRPFPFAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Qy	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP	180
Db	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP	180
Qy	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGES	300
Db	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGES	300
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Db	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Qy	361	PVREEYADFKPFEQAWVEVKDITYEGSRDVLAAARANVESKVDRKCIEDSLEQKSLGKDSEGR	420
Db	361	PVREEYADFKPFEQAWVEVKDITYEGSRDVLAAARANVESKVDRKCIEDSLEQKSLGKDSEGR	420

Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLLEPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLLEPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Qy	961	RSLSAVLSAELSKT	974
Db	961	RSLSAVLSAELSKT	974

RESULT 6

AAU04591

ID AAU04591 standard; protein; 1192 AA.

XX

AC AAU04591;

XX

DT 26-SEP-2001 (first entry)

XX

DE Human Nogo protein.

XX

KW Human; Nogo receptor; axonal growth; immunogen; antibody; nogo protein;

KW cranial trauma; cerebral trauma; spinal cord injury; stroke;

KW demyelinating disease; multiple sclerosis; monophasic demyelination;

KW encephalomyelitis; multifocal leukoencephalopathy; panencephalitis;

KW Marchiafava-Bignami disease; pontine myelinolysis; adrenoleukodystrophy;

KW Pelizaeus-Merzbacher disease; Spongy degeneration; Alexander's disease;

KW Canavan's disease; metachromatic leukodystrophy; viral infection;  
KW Krabbe's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 1054. .1119

FT /label= Lumenal\_extracellular\_domain

FT /note= "This sequence is specifically claimed"

FT Peptide 1055. .1094

FT /label= Pep1

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide 1064. .1088

FT /label= Pep2

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide 1074. .1098

FT /label= Pep3

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide 1084. .1108

FT /label= Pep4

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

FT Peptide 1095. .1119

FT /label= Pep5

FT /note= "Receptor binding inhibitory peptide. This  
FT sequence is specifically claimed"

XX

PN WO200151520-A2.

XX

PD 19-JUL-2001.

XX

PF 12-JAN-2001; 2001WO-US001041.

XX

PR 12-JAN-2000; 2000US-0175707P.

PR 26-MAY-2000; 2000US-0207366P.

PR 29-SEP-2000; 2000US-0236378P.

XX

PA (UYYA ) UNIV YALE.

XX

PI Strittmatter SM;

XX

DR WPI; 2001-442138/47.

DR N-PSDB; AAS09453.

XX

PT Novel Nogo receptor protein useful for identifying modulator of Nogo  
PT protein or Nogo receptor protein, which is useful for treating central  
PT nervous system disorders.

XX

PS Example 1; Page 101-104; 109pp; English.

XX

CC The sequence is the human Nogo protein, a 250kDa myelin-associated axon  
CC growth inhibitor. The invention relates to the use of the nogo receptor,  
CC nogo protein, their nucleic acids, vectors expressing them and antibodies  
CC against them, to isolate agents which block nogo receptor mediated axonal



CC growth. The agent is useful for treating a central nervous system  
CC disorder which is a result of cranial or cerebral trauma, spinal cord  
CC injury, stroke or a demyelinating disease selected from multiple  
CC sclerosis, monophasis demyelination, encephalomyelitis, multifocal  
CC leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease,  
CC pontine myelinolysis, adrenoleukodystrophy, Pelizaeus-Merzbacher disease,  
CC Spongy degeneration, Alexander's disease, Canavan's disease,  
CC metachromatic leukodystrophy, viral infection and Krabbe's disease

XX

SQ Sequence 1192 AA;

Query Match 75.3%; Score 4403.5; DB 4; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 1.3e-221;  
Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSILVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || |||:|:| ||||| |||||: ||: |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||| ||||| ||||| ||| || ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
      || ||||| ||||| |||| ||| |||||:|:|||||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy    226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
      ||||| ||||| ||||| | :|||:| :|||:|:| :| :||| |||||
Db    238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy    286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
      ||||| ||| |||:| | :|:|:|:|:|:| | : ||: || | |||
Db    298 SEMGSSFVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED 357

Qy    340 RVVSPEKTMDIFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV 395
      ||| || | ||| :||| |||:||||| |||: ||| ||| :|:
Db    358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERWVEVKDSKEDS-DMLAAGGKIESNL 416

Qy    396 ESKVDRKCLEDSELEQSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454
      |||||:| ||||| : |||| |:| ||||| :|| | ||||| | :|||
Db    417 ESKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA 476

Qy    455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513
      | |||| | ||||| |||||:||||:||| || ||||| |||| | ||| :
Db    477 TNIFPLLGDPTSENKTDEKKIEEKKQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy    514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL 573
      |||| |:||||| ||||| |||||:||||| :||||| |||
Db    537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL 596

Qy    574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSIKLEPEN 633
      |||||:||||| ||||| :|||||:||| |||| |:|:||| |||
```

Db 597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN 655  
 Qy 634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF 692  
 Db 656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDF 715  
 Qy 693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS 752  
 Db 716 SDYSEMAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIQDVPQKQDETVMVLKESLTETS 775  
 Qy 753 -ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF 809  
 Db 776 FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL 835  
 Qy 810 NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDEFPTFVSAKDDS-PKLAKEYTDLEVS 868  
 Db 836 STAVYSNDDLFIKSEAQIRETETFSDDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVS 895  
 Qy 869 DKSEIANIQSGADSLPCLLEPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV 926  
 Db 896 HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVLLLPDV 955  
 Qy 927 SALEPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK 986  
 Db 956 SALATQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK 1015  
 Qy 987 KTGTVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA 1046  
 Db 1016 KTGTVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA 1075  
 Qy 1047 YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1106  
 Db 1076 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1135  
 Qy 1107 NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 1136 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 7

ABP68600

ID ABP68600 standard; protein; 1192 AA.

XX

AC ABP68600;

XX

DT 14-JAN-2003 (first entry)

XX

DE Human pancreatic cancer expressed protein SEQ ID NO 71.

XX

KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;  
 KW cytostatic; tumour.

XX

OS Homo sapiens.

XX

PN WO200260317-A2.

XX

PD 08-AUG-2002.

XX  
PF 30-JAN-2002; 2002WO-US002781.  
XX  
PR 30-JAN-2001; 2001US-0265305P.  
PR 31-JAN-2001; 2001US-0265682P.  
PR 09-FEB-2001; 2001US-0267568P.  
PR 21-MAR-2001; 2001US-0278651P.  
PR 28-APR-2001; 2001US-0287112P.  
PR 16-MAY-2001; 2001US-0291631P.  
PR 12-JUL-2001; 2001US-0305484P.  
PR 20-AUG-2001; 2001US-0313999P.  
PR 27-NOV-2001; 2001US-0333626P.

XX  
PA (CORI-) CORIXA CORP.

XX  
PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;  
XX  
DR WPI; 2002-627435/67.  
DR N-PSDB; ABV94680.

XX  
PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for  
PT diagnosing, preventing and/or treating cancer, particularly pancreatic  
PT cancer.

XX  
PS Claim 2; SEQ ID NO 71; 300pp + Sequence Listing; English.

XX  
CC The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)  
CC complements of (a); (c) sequences consisting of at least 20 contiguous  
CC residues of (a); (d) sequences that hybridize to (a), under moderately  
CC stringent conditions; (e) sequences having at least 75% or 90% identity  
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer  
CC in a patient and compositions comprising polypeptides, polynucleotides,  
CC antibodies, fusion proteins, T cell populations and antigen presenting  
CC cells expressing the polypeptide are useful in treating pancreatic cancer  
CC and stimulating an immune response. The polynucleotides can be used as  
CC probes or primers for nucleic acid hybridisation, in the design and  
CC preparation of ribozyme molecules for inhibiting expression of the tumour  
CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
CC therapy. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 1192 AA;

Query Match 75.3%; Score 4403.5; DB 5; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 1.3e-221;  
Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

QY 1 MEDIDQSSILVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60  
|||:||||| ||||| ||||| ||||| ||||| |||||:||||| ||||| |||||  
Db 1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58  
  
QY 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
||||||| || ||| |||:||| :| ||||| ||||| |||||: ||| : |  
Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSVDETLFALPAASEPVIPISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS	237
Qy	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
Db	238	LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMGSSFKGSPKGESAILVENTKKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
Db	298	SEMGSSFVSVPKAESAVIVANPREEIIIVKNKDEEEKLVSNNIHNNQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFEQAWVEVKDITYEGSRDVLAA----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDKRCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA	476
Qy	455	ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTDTLS	513
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT	536
Qy	514	KVTEAAVSNMPEGLTPDLVQACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
Db	537	KVTEEVVANMPEGLTPDLVQACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN	633
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN	655
Qy	634	PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPSPDF	692
Db	656	PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIIACDLIKETKLSAEPAPDF	715
Qy	693	SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVS	752
Db	716	SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMVLVKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
Db	776	FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDFPPTFVSAKDDS-PKLAKEYTDLEVS	868
Db	836	STAVYSNDDLFIKSEAQIRETETFSDDSPIEIIIDFPPTLISSKTDSEFSKLAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLLEPCDLSEFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
Db	896	HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSKVLLLPDVP	955



CC The invention relates to a novel nucleic acid encoding a polypeptide  
 CC comprising amino acid residues 27-309 of a 473 amino acid sequence (P1,  
 CC human Nogo receptor (NgR) NTLRRCT domain), or residues 27-309 of P1 with  
 CC 1-20 conservative amino acid substitutions, and less than a complete CTS  
 CC domain, provided that a partial CTS domain, if present, consists of no  
 CC more than the first 39 consecutive residues. The nucleic acid of the  
 CC invention has neuroprotective activity. The polynucleotide may have a use  
 CC in gene therapy. The nucleic acid is useful for decreasing inhibition of  
 CC axonal growth by a central nervous system (CNS) neuron. The NgR  
 CC polypeptide or an agent inhibits the binding of Nogo to NgR or NgR-  
 CC dependent signal transduction in the central nervous system neuron may be  
 CC used in treating central nervous system disease, disorder or injury, e.g.  
 CC spinal cord injury. Expression of an NgR protein may be associated with  
 CC inhibition of axonal regeneration following cranial, cerebral or spinal  
 CC trauma, stroke or a demyelinating disease, such as multiple sclerosis,  
 CC monophasic demyelination, encephalomyelitis, multifocal  
 CC leukoencephalopathy, panencephalitis, or Krabbe's disease. The present  
 CC sequence is used in the exemplification of the invention  
 XX  
 SQ Sequence 1192 AA;

Query Match 75.3%; Score 4403.5; DB 6; Length 1192;  
 Best Local Similarity 75.9%; Pred. No. 1.3e-221;  
 Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

QY	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEELEVLERK	60
		:	
Db	1	MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK	58
QY	61	PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP	118
QY	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKR	178
QY	167	RG-SGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS	237
QY	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
QY	286	SEMGSSFSGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
Db	298	SEMGSSFVSVPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED	357
QY	340	RVVSPEKTMDFNEMQMSVAVPREYADFKPFEQAWVKDITYEGSRDLVLA---RANV	395
Db	358	EVVSSEKAKDSFNEKRVAEAPMREYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
QY	396	ESKVDKRCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA	476



DE Human MAGI polypeptide.  
 XX  
 KW MAGI protein; neuroendocrine-specific protein; neuropathy; human;  
 KW spinal injury; neuronal degeneration; neuromuscular disorder; cancer;  
 KW psychiatric disorder; developmental disorder; inflammatory disorder;  
 KW stroke; cytostatic; cerebroprotective; neuroprotective.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200005364-A1.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-GB002360.  
 XX  
 PR 22-JUL-1998; 98GB-00016024.  
 PR 19-JUL-1999; 99GB-00016898.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Michalovich D, Prinjha RK;  
 XX  
 DR WPI; 2000-182693/16.  
 DR N-PSDB; AAZ56886.  
 XX  
 PT Novel polypeptides related to neuroendocrine-specific proteins and  
 PT polynucleotides useful for diagnosis of various diseases and for  
 PT treatment of cancer and neurological disorders.  
 XX  
 PS Claim 2; Page 20-21; 35pp; English.  
 XX  
 CC The invention relates to human MAGI protein, which is similar to  
 CC neuroendocrine-specific protein. The MAGI protein can be expressed by  
 CC standard recombinant methodology. The MAGI polypeptides, polynucleotides  
 CC and antibodies are useful for treating diseases, including neuropathies,  
 CC spinal injury, neuronal degeneration, neuromuscular disorders,  
 CC psychiatric disorders and developmental disorders, cancer, stroke and  
 CC inflammatory disorders. The polynucleoitde is also useful for chromosome  
 CC localization and for tissue expression studies. The present sequence  
 CC represents the human MAGI protein  
 XX  
 SQ Sequence 1192 AA;

Query Match 75.2%; Score 4398.5; DB 3; Length 1192;  
 Best Local Similarity 75.9%; Pred. No. 2.3e-221;  
 Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

Qy 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60  
 |||:|||| ||||| ||||| ||||| ||||| |||||:||||| |||||  
 Db 1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDDE-EEEEEEEEDEDEDLEEEVLERK 58  
 Qy 61 PAAGLSAAVP--PAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
 ||||| || ||| |||: || : ||||| ||||| |||||: || : |  
 Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP 118  
 Qy 116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
 ||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSVDETLFALPAASEPVIIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS	237
Qy	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
Db	298	SEMGSSFSVSPKAESAIVANPREEIIVKNKDEEEKLVSNNIHNNQQLPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDKRCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA	476
Qy	455	ANTFPILLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS	513
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT	536
Qy	514	KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
Db	537	KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN	633
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN	655
Qy	634	PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPSDF	692
Db	656	PPPYEEAMSVSLKKVSGIKEEIKPENINAALQETEAPYISIIACDLIKETKLSAEPAPDF	715
Qy	693	SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS	752
Db	716	SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPEVPQKQDETVMVLKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
Db	776	FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDEFPTFVSAKDDS-PKLAKEYTDLEVS	868
Db	836	STAVYSNDDLFIKESQIRETETFSDDSSPIEIIDEFPTLISSKTDTSFKLAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLLPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
Db	896	HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVLLLPDPV	955
Qy	927	SALEPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	986
Db	956	SALATQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK	1015

Qy 987 KTG VVFGASLFLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1046  
 |||||  
 Db 1016 KTG VVFGASLFLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1075

Qy 1047 YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALF 1106  
 |||||  
 Db 1076 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVFTYVGALF 1135

Qy 1107 NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 |||||:|||||:|||||:|||||:  
 Db 1136 NGLTLLILALISLFSVPVIYERHQVQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 10

AAB82349

ID AAB82349 standard; protein; 1192 AA.

XX

AC AAB82349;

XX

DT 23-JUL-2001 (first entry)

XX

DE Human NOGO-A protein.

XX

KW NOGO-A; human; chromosome 2p21; neuropathy; spinal injury; brain injury;  
 KW stroke; neuronal degeneration; Alzheimer's disease; Parkinson's disease;  
 KW neuromuscular disorder; psychiatric disorder; developmental disorder;  
 KW neuroprotective; nootropic; neuroleptic; antiparkinsonian;  
 KW cerebroprotective; neuroleptic; diagnosis; therapy.

XX

OS Homo sapiens.

XX

PN WO200136631-A1.

XX

PD 25-MAY-2001.

XX

PF 14-NOV-2000; 2000WO-GB004345.

XX

PR 15-NOV-1999; 99GB-00026995.

PR 24-JAN-2000; 2000GB-00001550.

XX

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX

PI Michalovich D, Prinjha R;

XX

DR WPI; 2001-343822/36.

DR N-PSDB; AAF90324.

XX

PT New polypeptide designated NOGO-C is a splice variant of the human NOGO  
 PT gene and may be useful in the treatment of neural disorders including  
 PT Alzheimer's and Parkinson's diseases.

XX

PS Disclosure; Page 26-27; 25pp; English.

XX

CC The present sequence is that of human NOGO-A. NOGO-A is a previously  
 CC known splice variant of the human NOGO gene on chromosome 2p21. The  
 CC invention relates to a novel splice variant, NOGO-C (see AAB82348). It





XX  
 OS Homo sapiens.  
 XX  
 PN WO200257483-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 18-JAN-2002; 2002WO-GB000228.  
 XX  
 PR 18-JAN-2001; 2001GB-00001312.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Blackstock WP, Hale RS, Prinjha R, Rowley A;  
 XX  
 DR WPI; 2002-599722/64.  
 DR N-PSDB; ABK90134.  
 XX  
 PT Identifying modulators of Nogo or BACE activity for treating acute  
 PT neuronal injuries, neoplastic or dysproliferative disorders, comprises  
 PT providing and monitoring interaction between Nogo and BACE polypeptides.  
 XX  
 PS Disclosure; Page 59-62; 68pp; English.  
 XX  
 CC The present invention relates to a new method of identifying modulators  
 CC of Nogo function or BACE activity. The method involves providing Nogo and  
 CC BACE polypeptides capable of binding with each other, monitoring the  
 CC interaction between these polypeptides, and determining if the test agent  
 CC is a modulator of Nogo or BACE activity. The method is useful in treating  
 CC acute neuronal injuries, such as spinal or head injury, stroke,  
 CC peripheral nerve damage, and in neoplastic (e.g. glioblastomas,  
 CC neuroblastomas), hyperproliferative or dysproliferative disorders (e.g.  
 CC cirrhosis, psoriasis, keloid formation, fibrocystic conditions, tissue  
 CC hypertrophy) of the central nervous system. The BACE polypeptide is  
 CC useful in screening methods to identify agents that may act as modulators  
 CC of BACE activity and in particular agents that may be useful in treating  
 CC Nogo-associated diseases. The modulators of Nogo or BACE polypeptides,  
 CC and the polynucleotide encoding the BACE polypeptide are useful in  
 CC manufacturing a medicament for the treatment or prevention of disorders  
 CC responsive to the modulation of Nogo activity, in alleviating the  
 CC symptoms or improving the condition of a patient suffering from this  
 CC disorder, in axon regeneration, or in preventing metastasis or spreading  
 CC of a cancer. The polynucleotide may also be an essential component in  
 CC assays, a probe, in recombinant protein synthesis, and in gene therapy  
 CC techniques. The present amino acid sequence represents the human NogoA  
 CC protein of the invention  
 XX  
 SQ Sequence 1192 AA;

Query Match 75.2%; Score 4398.5; DB 5; Length 1192;  
 Best Local Similarity 75.9%; Pred. No. 2.3e-221;  
 Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

Qy 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60  
 |||:||||| ||||| ||||| ||||| ||||| |||||:|:|:|||||  
 Db 1 MEDLDQSPVSSS-DSPPRQPAPFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58

Qy 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
 ||||| || ||| ||:| :| ||||| ||||| : || : |  
 Db 59 PAAGLSAAPVFTAPAAAGAPLMDFGNDFFVPPAPRGPLPAAPPVAPERQPSWDPSFVSSTVP 118

Qy 116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
 ||| ||||| ||||| ||||| ||||| || || ||||| |||||  
 Db 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy 167 RG-SGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225  
 || ||||| ||||| ||||| ||||| ||||| :||| ||||| |||||  
 Db 179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy 226 LSPLSTVSFKEHGYLGNL SAVSSSEG TIEETLNEASKELPERATNPFVNRDLAEFSELEY 285  
 ||||| ||||| ||||| | :|||:| :|||:| :| :||| |||||  
 Db 238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy 286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339  
 ||||| ||| |||:| | :||:|:|:|:|:| | : ||: || | |||  
 Db 298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED 357

Qy 340 RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV 395  
 ||| || | ||| :||| ||:|||||||: |||||: | | |:||| :||  
 Db 358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 416

Qy 396 ESKVDRKCLEDSELEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454  
 |||||:| | |||| : |||| |:|| ||||| :|| ||||| | :|||  
 Db 417 ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA 476

Qy 455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513  
 | |||| | ||||| |||||:||||:| || ||||| |||| | ||||| |:  
 Db 477 TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy 514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL 573  
 |||| |:||||| ||||| ||||| |||||:||||| :|||| |||  
 Db 537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL 596

Qy 574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN 633  
 |||||:||||| ||||| :|||||:||| |||| |:|||| ||||  
 Db 597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN 655

Qy 634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF 692  
 |||||:|:| | : | || ||||: |||:||||| ||||| ||:|  
 Db 656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDF 715

Qy 693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS 752  
 |:|:|:| | : ||:|:|||||:||||| |||||:| | | ||:| ||||| |  
 Db 716 SDYSEMAKVEQVPDHSSELVEDSSPDSEPVDLFSDDSIPEVPQKQDETVM LVKESLTETS 775

Qy 753 -ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF 809  
 |:: :: :|:| | | |||||: :| :|| :: :|:| |||||  
 Db 776 FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL 835

Qy 810 NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPTFVSAKDDS-PKLAKEYTDLEVS 868  
 :||:| ||| :|:|:||||| ||||| :| | |||:|||||  
 Db 836 STAVYSNDDLFIKAEQIRETETFSDDSSPIEIIIDEFPTLISSKTD SFSKLAREYTDLEVS 895



DR N-PSDB; ABN86601.

XX

PT Promoting nerve regeneration and preventing neuronal degeneration in the  
PT central/peripheral nervous system from injury/disease, comprises  
PT administering nervous system-specific activated T cells/antigen, or  
PT analogs/peptides.

XX

PS Example; Page 53-56; 93pp; English.

XX

CC The invention relates to promoting nerve regeneration or conferring  
CC neuroprotection and preventing or inhibiting neuronal degeneration in the  
CC central/peripheral nervous system (NS). The method involves administering  
CC NS-specific activated T cells, NS-specific antigen, its analogue or its  
CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or  
CC combinations. The method is useful for promoting nerve regeneration and  
CC preventing neuronal degeneration in central/peripheral nervous system  
CC from injury/disease, where the injury is spinal cord injury, blunt  
CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or  
CC damages caused by surgery such as tumour excision. The disease is not an  
CC autoimmune disease or neoplasm. The disease results in a degenerative  
CC process occurring in either gray or white matter or both. The disease is  
CC diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's  
CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,  
CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and  
CC vitamin deficiency, intervertebral disc herniation, prion diseases such  
CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral  
CC neuropathies associated with various diseases, including but not limited  
CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute  
CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary  
CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption  
CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-  
CC pathies, complications of various drugs (e.g., metronidazole) and toxins  
CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia  
CC telangectasia, Friedreich's ataxia, amyloid polyneuropathies,  
CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's  
CC disease, or lipoproteinemia. The present sequence represents the human  
CC neurotransmitter receptor protein Nogo-A, an example of NS-specific  
CC antigen

XX

SQ Sequence 1192 AA;

Query Match 75.2%; Score 4398.5; DB 5; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 2.3e-221;  
Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |||:||||| ||||| ||||| ||||| ||||| |||||:||||| |||||
Db      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEDEDEDLEEEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || ||| |||:| :| ||||| ||||| |||||: || : |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPPAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| || || ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPSTPAAPKR 178
```



Qy 167 RG-SGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225  
 || ||||| ||||| ||||| ||||| :||| ||||| ||||| |||||

Db 179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy 226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285  
 ||||| ||||| ||||| ||||| :||| :||| :||| :||| :||| :||| :|||

Db 238 LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy 286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339  
 ||||| ||||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED 357

Qy 340 RVVSPEKTMDIFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA-----RANV 395  
 ||||| ||||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 416

Qy 396 ESKVDRKCLEDSELEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454  
 ||||| :||| ||||| :||| ||||| :||| ||||| :||| |||||

Db 417 ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA 476

Qy 455 ANTFFLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFVLVAVQDSEADYVTTDTLS 513  
 | |||| | ||||| :||| :||| :||| :||| :||| :||| :|||

Db 477 TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFVLVAAQDSETDYVTTDNL 536

Qy 514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTS EAIQESLYPTAQL 573  
 |||| | :||| :||| :||| :||| :||| :||| :||| :|||

Db 537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTS EVMQESLYPAAQL 596

Qy 574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN 633  
 ||||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN 655

Qy 634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQTEAPYISIACDLIKETKLSTEPSPDF 692  
 ||||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQTEAPYISIACDLIKETKLSAEPAPDF 715

Qy 693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDFSDDSIPEVPQTQEEAVMLMKESLTVS 752  
 | :||| :||| | :||| :||| :||| :||| :||| :||| :|||

Db 716 SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMVLVKESLTETS 775

Qy 753 -ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF 809  
 | :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 776 FESMIEYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL 835

Qy 810 NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDEFPTFVSAKDDS-PKLAKEYTDLEVS 868  
 :||| :||| | :||| :||| :||| :||| :||| :||| :|||

Db 836 STAVYSNDDLFIKAEQIRETETFSDDSSPIEIIDEFPTLISSKTDSEFSLAREYTDLEVS 895

Qy 869 DKSEIANIQSGADSLPCLELPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV 926  
 ||||| ||||| ||||| ||||| ||||| ||||| :||| :||| :||| :|||

Db 896 HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNKSATSKVLLLPDPV 955

Qy 927 SALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK 986  
 |||| | :||| | :||| :||| :||| :||| :||| :||| :|||

Db 956 SALATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK 1015

Qy 987 KTGTVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA 1046

```

Db      1016 KTG VVFGASL FLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1075
Qy      1047 YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVD SLKFAVLMWVF TYVGALF 1106
Db      1076 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVD SLKFAVLMWVF TYVGALF 1135
Qy      1107 NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163
Db      1136 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

```

RESULT 13

AAY71311

ID AAY71311 standard; protein; 1178 AA.

XX

AC AAY71311;

XX

DT 02-NOV-2000 (first entry)

XX

DE Human neurite growth inhibitor Nogo.

XX

KW Human; neurite growth inhibitor; Nogo; neural cell; myelin; CNS;  
KW central nervous system; neoplastic disease; antiproliferative; glioma;  
KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
KW structural plasticity; screening.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 187

FT /label= Unknown

FT Misc-difference 188

FT /label= Unknown

FT Misc-difference 189

FT /label= Unknown

FT Misc-difference 190

FT /label= Unknown

FT Misc-difference 221

FT /label= Unknown

FT Misc-difference 328

FT /label= Unknown

FT Misc-difference 477

FT /label= Unknown

FT Region 977. .1012

FT /note= "Region specifically described in claim 16"

FT Region 994. .1174

FT /note= "Region specifically described in claim 16"

FT Region 1079. .1114

FT /note= "Region specifically described in claim 16"

XX

PN WO200031235-A2.

XX

PD 02-JUN-2000.

XX  
 PF 05-NOV-1999; 99WO-US026160.  
 XX  
 PR 06-NOV-1998; 98US-0107446P.  
 XX  
 PA (SCHW/) SCHWAB M E.  
 PA (CHEN/) CHEN M S.  
 XX  
 PI Schwab ME, Chen MS;  
 XX  
 DR WPI; 2000-400052/34.  
 XX  
 PT Nogo proteins and nucleic acids useful for treating neoplastic disorders  
 PT of the central nervous system and inducing regeneration of neurons.  
 XX  
 PS Claim 11; Fig 13; 122pp; English.  
 XX  
 CC The present sequence is a human Nogo protein which is a potent neural  
 CC cell growth inhibitor and is free of all central nervous system (CNS)  
 CC myelin material with which it is natively associated. The human Nogo  
 CC sequence was derived by aligning human expressed sequence tags (ESTs)  
 CC e.g. AA158636, AA333267, AA081783, AA167765, AA322918, AA092565, AA081525  
 CC and AA081840 with the rat Nogo sequence. Nogo proteins and fragments  
 CC displaying neurite growth inhibitory activity are used in the treatment  
 CC of neoplastic disease of the CNS e.g. glioma, glioblastoma,  
 CC medulloblastoma, craniopharyngioma, ependyoma, pinealoma,  
 CC haemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma,  
 CC neuroblastoma or retinoblastoma and degenerative nerve diseases e.g.  
 CC Alzheimer's and Parkinson's diseases. Therapeutics which promote Nogo  
 CC activity can be used to treat or prevent hyperproliferative or benign  
 CC dysproliferative disorders e.g. psoriasis and tissue hypertrophy.  
 CC Ribozymes or antisense Nogo nucleic acids can be used to inhibit  
 CC production of Nogo protein to induce regeneration of neurons or to  
 CC promote structural plasticity of the CNS in disorders where neurite  
 CC growth, regeneration or maintenance are deficient or desired. The animal  
 CC models can be used in diagnostic and screening methods for predisposition  
 CC to disorders and to screen for or test molecules which can treat or  
 CC prevent disorders or diseases of the CNS. Note: SEQ ID numbers 35-42 are  
 CC referred in claim 32 and SEQ ID NO: 29 in disclosure of the  
 CC specification. However the specification does not include sequences for  
 CC these SEQ ID numbers  
 XX  
 SQ Sequence 1178 AA;

Query Match 73.1%; Score 4276.5; DB 3; Length 1178;  
 Best Local Similarity 73.8%; Pred. No. 5.5e-215;  
 Matches 883; Conservative 104; Mismatches 158; Indels 51; Gaps 19;

QY 1 MEDIDQSSIVSSSTDSPRPPPAFKYQFVTEPEDEDEDEEEDEEDEDLEEELEVLERK 60  
 |||:||||| || ||| ||||| ||||| |||||:|:|:|||||  
 Db 1 MEDLDQSPLVSSS-DSVPRPQPAFKYQFVREPEDEE-EEEEEEDEDEDEDLEEELEVLERK 58  
 QY 61 PAAGLSAAAVP---PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
 ||||| || ||| |||:|:| ||||| |||||: || : |  
 Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP 118  
 QY 116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166

Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RGSGSVDETLFALPAASEPVIIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL	226
Db	179	RGSSG-----AVVXXXXKIMDLKEQPGNTISAGQEDFPSVLLETAASXPSSL	224
Qy	227	SPLSTVSFKEHGYLGNL SAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYS	286
Db	225	SPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYS	284
Qy	287	EMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KEDR	340
Db	285	EMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHXQQELPTALTCLVKEDE	344
Qy	341	VVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANVE	396
Db	345	VVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNLE	403
Qy	397	SKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTTA	455
Db	404	SKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIAT	463
Qy	456	NTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLSK	514
Db	464	NIFPLLEDPTSENXTDEKKIEEKKAQIVTEKNTSTKTSNPFFVAAQDSCTDYVTTDNLT	523
Qy	515	VTEAAVSNMPEGLTPDLVQFACESELNEATGTKIAYETKVDLVQTSIAIQESLYPTAQLC	574
Db	524	VTEEVVANMPEGLTPDLVQFACESELNEVTGTKIAYETKMDLVQTSSEVMQESLYPAAQLC	583
Qy	575	PSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENP	634
Db	584	PSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPENP	642
Qy	635	PPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSN	694
Db	643	PPYEEAMSVSLKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSD	702
Qy	695	YSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS-E	753
Db	703	YSEMAKVEQVPVDHSELVEDSSPDSEPVDLFSDDSIPEVDPQKQDETVMVLKESLTETSFE	762
Qy	754	TVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEFNT	811
Db	763	SMIEYENKEKLSALPPEGGKPYLESFKLSLVNTKDTLLPDEVSTLSKKEKIPLQMEELST	822
Qy	812	AIYSNDDLSSKEDKIKESETFSDSSPIEIDEFPTFVSAKDDS-PKLAKEYTDLEVSDK	870
Db	823	AVYSNDDLFIKSEAQIRETETFSDSSPIEIDEFPTLISSKTDIFSGLAREYTDLEVSHK	882
Qy	871	SEIANIQSGADSLPCLELPCDLSEFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNVSA	928
Db	883	SEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNKSATSQVLLLPDVS	942
Qy	929	L-EPQTEMGSIIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	987

Db 943 LGHTQAEIESIVKPKVLEKEAEKKLPDTEKEDRSPSAIFSADLGKTSVVDLLYWRDIKK 1002  
 Qy 988 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAY 1047  
 Db 1003 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPPFRAY 1062  
 Qy 1048 LESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALEN 1107  
 Db 1063 LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALEN 1122  
 Qy 1108 GLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 1123 GLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1178

RESULT 14

AAU33228

ID AAU33228 standard; protein; 1246 AA.

XX

AC AAU33228;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #3719.

XX

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US008656.

XX

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-611725/70.

XX

PT Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy.

XX

PS Claim 20; Page 737; 765pp; English.

XX

CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 1246 AA;

Query Match 70.4%; Score 4116; DB 4; Length 1246;  
 Best Local Similarity 72.0%; Pred. No. 1.4e-206;  
 Matches 873; Conservative 112; Mismatches 171; Indels 56; Gaps 27;

Qy 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60  
 |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||  
 Db 42 MEDLDQSPLVSSS-DSPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 99  
 Qy 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
 ||||| || ||| |||:| | ||||| ||||| |||||: || : |  
 Db 100 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSFVSSTVP 159  
 Qy 116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
 ||| |||| ||||| ||||| ||||| || || |||||  
 Db 160 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 219  
 Qy 167 RG-SGSVDETLFALPAASEFVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225  
 || ||||| ||||| ||||| ||| |||||:|:|||||  
 Db 220 RGSSGSVDETLFALPAASEFVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 278  
 Qy 226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285  
 ||||| ||||| ||||| | :|||:| :|||: |:| :||| |||||  
 Db 279 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 338  
 Qy 286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESVPG-----KED 339  
 ||||| || |||:| | :||:|:|:|:| | : ||: || | |||  
 Db 339 SEMGSSFSVSPKAESAVIVANPREETIVKNKDEEEKLVSNILHNQQELPTALTKLVKED 398  
 Qy 340 RVVSPEKTMDFNEMQMSVVAPVREYADFKPFEQAWEVKDTYEGSRDVLAA-----RANV 395  
 ||| || | ||| :|| ||:|||||||: ||||: | | |:||| :|:  
 Db 399 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 457  
 Qy 396 ESKVDRKCLEDSELEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454  
 ||||:| |||| : |||| |:| ||||| :|| | ||||| | :|||  
 Db 458 ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA 517  
 Qy 455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513  
 | |||| | ||||| |||||:||||:|| | ||||| ||||| ||||| |:  
 Db 518 TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 577  
 Qy 514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSIAIQESLYPTAQL 573  
 |||| |:||||| ||||| |||||:||||| :||||| |||  
 Db 578 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSSEVMQESLYPAAQL 637



KW central nervous system; neoplastic disease; antiproliferative; glioma;  
KW antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;  
KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease;  
KW hyperproliferative disorder; benign dysproliferative disorder; diagnosis;  
KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment;  
KW structural plasticity; screening; mutant; mutein.

XX

OS Rattus sp.

XX

PN WO200031235-A2.

XX

PD 02-JUN-2000.

XX

PF 05-NOV-1999; 99WO-US026160.

XX

PR 06-NOV-1998; 98US-0107446P.

XX

PA (SCHW/) SCHWAB M E.

PA (CHEN/) CHEN M S.

XX

PI Schwab ME, Chen MS;

XX

DR WPI; 2000-400052/34.

XX

PT Nogo proteins and nucleic acids useful for treating neoplastic disorders  
PT of the central nervous system and inducing regeneration of neurons.

XX

PS Example; Page; 122pp; English.

XX

CC The patent relates to neurite growth inhibitor Nogo which is free of all  
CC central nervous system (CNS) myelin material with which it is natively  
CC associated. Nogo proteins and fragments displaying neurite growth  
CC inhibitory activity are used in the treatment of neoplastic disease of  
CC the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,  
CC ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,  
CC oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and  
CC degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.  
CC Therapeutics which promote Nogo activity can be used to treat or prevent  
CC hyperproliferative or benign dysproliferative disorders e.g. psoriasis  
CC and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be  
CC used to inhibit production of Nogo protein to induce regeneration of  
CC neurons or to promote structural plasticity of the CNS in disorders where  
CC neurite growth, regeneration or maintenance are deficient or desired. The  
CC animal models can be used in diagnostic and screening methods for  
CC predisposition to disorders and to screen for or test molecules which can  
CC treat or prevent disorders or diseases of the CNS. The present sequence  
CC is a fragment of rat Nogo A protein shown in AAY71310, which is used in  
CC the construction of mutant NiG. The mutant is composed of His-tag/T7-  
CC tag/Nogo-A sequence aa 172-974/His-tag. Nogo A deletion mutants were used  
CC for mapping the inhibitory sites of Nogo protein. Major inhibitory region  
CC was identified in the Nogo A sequence from amino acids 172-974,  
CC particularly amino acids 542-722. In addition, N-terminal region 1-171  
CC was found to be inhibitory to NIH 3T3 fibroblast spreading. Note: The  
CC present sequence is not given in the specification but is derived from  
CC rat Nogo A sequence shown in AAY71310. SEQ ID numbers 35-42 are referred  
CC in claim 32 and SEQ ID NO: 29 in disclosure of the specification.  
CC However, the specification does not include sequences for these SEQ ID



CC numbers

XX

SQ Sequence 803 AA;

Query Match 68.8%; Score 4023; DB 3; Length 803;

Best Local Similarity 99.9%; Pred. No. 6e-202;

Matches 802; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      172 VDETLFALPAASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLST 231
          |||
Db        1 VDETLFALPAASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLST 60

Qy      232 VSFKEHGYLGNL SAVSSSEGTIEETLINEASKELPERATNPFVNRDLAEFSELEYSEMGSS 291
          |||
Db        61 VSFKEHGYLGNL SAVSSSEGTIEETLINEASKELPERATNPFVNRDLAEFSELEYSEMGSS 120

Qy      292 FKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDF 351
          |||
Db       121 FKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDF 180

Qy      352 NEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSELEQK 411
          |||
Db       181 NEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSELEQK 240

Qy      412 SLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTD 471
          |||
Db       241 SLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTD 300

Qy      472 EKKIEERKAQIITEKTSNPFVLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDL 531
          |||
Db       301 EKKIEERKAQIITEKTSNPFVLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDL 360

Qy      532 VQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDI 591
          |||
Db       361 VQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDI 420

Qy      592 VMEAPLNSLLPSAGASVVQPSVSPLEAPPVSYDSIKLEPENPPPYEEAMNVALKALGTK 651
          |||
Db       421 VMEAPLNSLLPSAGASVVQPSVSPLEAPPVSYDSIKLEPENPPPYEEAMNVALKALGTK 480

Qy      652 EGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAEL 711
          |||
Db       481 EGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAEL 540

Qy      712 VEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELG 771
          |||
Db       541 VEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELG 600

Qy      772 KPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNNTAIYSNDDLSSKEDKIKESE 831
          |||
Db       601 KPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNNTAIYSNDDLSSKEDKIKESE 660

Qy      832 TFS DSSPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCD 891
          |||
Db       661 TFS DSSPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCD 720

Qy      892 LSFKN IY PKDEVHVSDEFSEN RSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKK 951
```

```

Db          |||||||||||||||||||
721 LSFKN IY PKDEVHVSDEFSEN RSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKK 780

Qy          952 LPSDTEKEDRSLSAVLSAELSKT 974
          |||||||||||||||
Db          781 LPSDTEKEDRSLSAVLSAELSKT 803

```

Search completed: September 29, 2004, 18:13:54  
Job time : 117.067 secs

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:09:33 ; Search time 28.8717 Seconds  
(without alignments)  
2079.581 Million cell updates/sec

Title: US-09-830-972-2  
Perfect score: 5848  
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRRKAD 1163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

8						
Result	Query			Description		
No.	Score	Match	Length	DB	ID	
1	908	15.5	199	2	US-08-700-607-1	Sequence 1, Appli
2	789.5	13.5	776	2	US-08-700-607-5	Sequence 5, Appli
3	716.5	12.3	356	2	US-08-700-607-6	Sequence 6, Appli
4	685	11.7	208	2	US-08-700-607-7	Sequence 7, Appli
5	671	11.5	267	2	US-08-700-607-8	Sequence 8, Appli
6	539.5	9.2	168	4	US-09-149-476-563	Sequence 563, App
7	519	8.9	241	2	US-08-700-607-3	Sequence 3, Appli
8	302.5	5.2	8991	4	US-08-714-741-32	Sequence 32, Appl
9	288	4.9	2468	4	US-09-976-594-726	Sequence 726, App
10	286	4.9	92	4	US-09-149-476-411	Sequence 411, App
11	279.5	4.8	1786	3	US-08-973-462-8	Sequence 8, Appli

12	265.5	4.5	1780	1	US-08-769-309A-5	Sequence 5, Appli
13	265.5	4.5	1780	3	US-08-994-570-5	Sequence 5, Appli
14	265	4.5	2137	4	US-09-134-001C-4463	Sequence 4463, Ap
15	256.5	4.4	75	4	US-09-621-976-4600	Sequence 4600, Ap
16	256.5	4.4	75	4	US-09-621-976-4601	Sequence 4601, Ap
17	254.5	4.4	1596	4	US-08-978-277A-4	Sequence 4, Appli
18	233	4.0	1142	2	US-08-993-118-7	Sequence 7, Appli
19	233	4.0	1142	3	US-08-845-528C-7	Sequence 7, Appli
20	233	4.0	1142	4	US-09-066-281B-7	Sequence 7, Appli
21	233	4.0	1142	4	US-09-468-433C-7	Sequence 7, Appli
22	232.5	4.0	1805	1	US-07-853-913-2	Sequence 2, Appli
23	230	3.9	3256	4	US-09-919-172-98	Sequence 98, Appl
24	230	3.9	3256	4	US-09-976-594-22	Sequence 22, Appl
25	228	3.9	1142	3	US-09-061-709-2	Sequence 2, Appli
26	228	3.9	1142	4	US-09-899-651-2	Sequence 2, Appli
27	228	3.9	1142	4	US-09-392-714-26	Sequence 26, Appl
28	227	3.9	1270	4	US-07-757-022B-44	Sequence 44, Appl
29	227	3.9	1311	4	US-07-757-022B-42	Sequence 42, Appl
30	227	3.9	1313	4	US-07-757-022B-142	Sequence 142, App
31	227	3.9	1314	4	US-07-757-022B-50	Sequence 50, Appl
32	227	3.9	1320	4	US-07-757-022B-46	Sequence 46, Appl
33	227	3.9	1320	4	US-07-757-022B-60	Sequence 60, Appl
34	227	3.9	1354	4	US-07-757-022B-48	Sequence 48, Appl
35	227	3.9	1361	4	US-07-757-022B-40	Sequence 40, Appl
36	227	3.9	1363	4	US-07-757-022B-52	Sequence 52, Appl
37	227	3.9	1404	4	US-07-757-022B-2	Sequence 2, Appli
38	227	3.9	1404	4	US-07-757-022B-62	Sequence 62, Appl
39	225	3.8	688	3	US-09-141-047-8	Sequence 8, Appli
40	224.5	3.8	941	4	US-07-757-022B-14	Sequence 14, Appl
41	224.5	3.8	1022	4	US-07-757-022B-84	Sequence 84, Appl
42	224.5	3.8	1038	4	US-07-757-022B-74	Sequence 74, Appl
43	224.5	3.8	1049	4	US-07-757-022B-58	Sequence 58, Appl
44	224.5	3.8	1140	4	US-07-757-022B-104	Sequence 104, App
45	224.5	3.8	1346	2	US-08-635-121-2	Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-08-700-607-1

; Sequence 1, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304



```

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307307
US-08-700-607-5

```

```

Query Match          13.5%; Score 789.5; DB 2; Length 776;
Best Local Similarity 31.2%; Pred. No. 3.1e-42;
Matches 243; Conservative 115; Mismatches 252; Indels 169; Gaps 28;

```

```

Qy      487 TSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEQACESELNEATGT 546
       : |  : |  ||: :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      65 SGPARGSP--VAMETASTGVAGVSSAMDHTFSTTSKDGE-----SCYTSLI----S 110

Qy      547 KIAYETKVDLVQTSEAIQ-ESLYPTAQLCPSFEEAEATPSPVLPDI--VMEAPLNS---- 599
       | |  : |  :  : | : : |  |  |  |  |  |  |  :  :  :  |  |
Db     111 DICYPQEDSTYFTGILQKENGHTISESP---EELGTPGPSLPDVPGLIESRGLFSSDSG 167

Qy      600 --LLPSAGASVVQPSVSPLEAPPPVSY-----DSIKLEPENPPPYEEA-----M 641
       : | :  | :  ||: :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     168 IEMTPAESTEVNKLADPLDQMKA EAYKYIDITRPEEVKHQEQHHPELEDKDLDFKNKDT 227

Qy      642 NVALKALGTKEGIKEPE-----SFNAAVQETEAPYISIACDLIKETKLSTE-PSP 690
       : : : |  ||: ||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     228 DISIK----PEGVREPDKPAPVEGKIIKDHLLEESTFAPYID---DLSEEQRRAPQITTP 280

Qy      691 DFSNYSEIAKFEKSVPEHAELVEDSSPESEPVDLF-SDDSIPEV----PQTQEEAVMLMK 745

```

```

      :||      |      :: :|| : : | | :|| |      | :      :
Db      281 VKITLTEIE-----PSVETTTQEKTPKQDICKPSPDTVPTVTVSEPEDDSPGSIPTP 334
Qy      746 ESLTEVSETVAQHK-----EERLSASPQELGKPYLESFQP---NLHSTKDAASND---- 792
      | || | :| | :| :| : | | : | :      :
Db      335 SSGTEPSAAESQGKGSISEDELITAIKEAKGLSYETAENPRPVGQLADRPEVKARSGPPT 394
Qy      793 IPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKI----- 827
      ||: | | : : : | | : : || :
Db      395 IPSPLDHEASSAESGDSEIELVSEDPM--AAEDALPSGYVSFGHVGGPPPPSPASPSIQYS 452
Qy      828 ---KESETFSDSSPIEIIDFPTFVSAKDDSPKLAKEYTDLEVSDKSEI-----ANIQSG 879
      :| | || |:| | || :||| : : :| | :
Db      453 ILREEREAElds---ELIIESCDASSASEESPKREQDSPMKPSALDAIREETGVRAEER 509
Qy      880 ADSLPCLELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSI 939
      | | | | | : || | | | | | | | | | | |
Db      510 APSRRGLAEPG--SFLD-YPSTEPQPGPEL-----PPGDGALEPETPM---- 549
Qy      940 KSKSLTKEAEKKLPSDTEKEDRSLSAVLSAE-----LSKTSVVDLLYWRD 984
      || |:| : : : : | :| :||| | |
Db      550 -----LPRKPEEDSSSNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLlyWRD 597
Qy      985 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPF 1044
      ||:|:|:|:| | || | |:| | |:| | | | | | | | | | | | | | | | | | | | | |
Db      598 IKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPF 657
Qy      1045 RAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGA 1104
      :||| | : :|:| :||:| :|||:| | | | | | | | | | | | | | | | | | | | | |
Db      658 KAYLELEITLSQEIQKYTDCLQFYVNSTLKLRLFLVQDLVDLSLKFAVLMWLLTYVGA 717
Qy      1105 LFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163
      | | | | | | | :|:|:|:|:|:| | :| | | | | | : :| | | | | | | | | |
Db      718 LFNGLTLLMAVVSMTFLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHAE 776

```

RESULT 3

US-08-700-607-6

; Sequence 6, Application US/08700607

; Patent No. 5858708

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Goli, Surya K.

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible





```

; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307311
US-08-700-607-7

```

```

Query Match          11.7%; Score 685; DB 2; Length 208;
Best Local Similarity 67.0%; Pred. No. 2.2e-36;
Matches 128; Conservative 32; Mismatches 31; Indels 0; Gaps 0;

```

```

Qy      973 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGI 1032
      |: :|||||||:|:|:|: | || || | |:||| |:||| || ||||| |:
Db      18 KSQAIDLLYWRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALAALSATISFRIYKSVL 77

Qy     1033 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDILVDSLKF 1092
      ||:|:|:|||||:|:| | |: :|:| :|||: :|||:|:||||| | |||||
Db      78 QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQFYVNSTLKLRLFLVQDLVDSLKF 137

Qy     1093 AVLMMWFVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ 1152
      ||||: | ||||| |||||:|:|:|:|:|:|:| :| || | ||| : :|||
Db     138 AVLMWLLTYVGALFNGLTLLLMVSMFTLPVVYVKHQAQIDQYLGIVRTHINAVVAKIQ 197

```





; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594

```
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
```

```
Query Match          9.2%;  Score 539.5;  DB 4;  Length 168;
Best Local Similarity 59.3%;  Pred. No. 3.2e-27;
Matches 99;  Conservative 36;  Mismatches 31;  Indels 1;  Gaps 1;
```

```
Qy      998 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEE 1057
      :|||  ||::||  ::  |||||  |||||  |||||  |||||  ::  :|  |
Db      1 MLLSLAAFSVISVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSE 60

Qy     1058 LVQKYSNSALGHVNSTIKELRRFLVDDLDVDSLKFAVLMWVFTYVGALFNGLTLLILALI 1117
      |  |::  |::  :|  :  |||||  |||||  ||  ||  |||||  |||||  |||||  :
Db      61 AFHNYMNAAMVHINRALKLIIRLFLVEDLDVDSLKLAVFMWLMFTYVGAVFNGITLLILAE 120
```



Db 44 SSCAVHDLIXWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSV 103

Qy 1032 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRFLVDDLVDLSLK 1091  
 |||:||||:||||:||||: :: :| | | |:|: |:| :| : |||||:|||||||

Db 104 IQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHINRALKLIIRLFLVEDLVDLSLK 163

Qy 1092 FAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQV 1131  
 || ||: |||||:||||:||||||| : : |:|:| :::|

Db 164 LAVFMWLMTYVGAVFNGITLLILAELLIXSVPIVYXKYKV 203

RESULT 8

US-08-714-741-32

; Sequence 32, Application US/08714741

; Patent No. 6500613

; GENERAL INFORMATION:

; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Crain, Marilyn J.

; APPLICANT: Hollingshead, Susan

; APPLICANT: Tart, Rebecca

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,

; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,

; TITLE OF INVENTION: PORTIONS AND PRODUCTS

; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: New York

; COUNTRY: U.S.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/714,741

; FILING DATE: 16-SEP-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer Esq., William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454312-2460

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 32:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8991 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear



; MOLECULE TYPE: amino acid  
US-08-714-741-32

Query Match 5.2%; Score 302.5; DB 4; Length 8991;  
Best Local Similarity 22.1%; Pred. No. 1.6e-09;  
Matches 244; Conservative 125; Mismatches 439; Indels 297; Gaps 47;

```
Qy      18 PRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERKPAAGLSAAAVPPAAAAAP 77
      |:| || :      :|| :||: ||| | || | :|      || | |
Db      7805 PKPAPAPQPAPAPKPEKTDQQAEEEDYARRSE--EEYNRLPQQQ-----PPKAEKP 7853

Qy      78 LLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLPPAAAVLP SKLPEDDEPP 137
      : | | | :|| | | | | | | | | | | | | | | |
Db      7854 -----APAPKPEQPVPAEXPENP-----APAPK--PAXAPQPLKPEEPAEQP 7893

Qy     138 ARPPPPPPAGASPLAEPAPAPPSTPAAP--KRRGSGSVDETFLFALPAASEPVI PSSAEKIM 195
      | ||| | | | | | | | : | : | |
Db      7894 KPEKPEEPAGQ---PEPEKPDQQAAGEDYARRSGGEYNRFPQQQPPKAEK--PAPAPK-- 7946

Qy     196 DLMEQPGNTVSSGQEDFSPVLLETAASLPSLSPLSTVSF-----KEHGYLG NLSAVSS 248
      ||| | | : | : | : | | | | :
Db      7947 --PEQP-----VPAPKTLLKKAKLAGAKSKAATKKAELEPELEKAEAELENLLSTLD 7996

Qy     249 SEGTTIETLNEASKELPERATNPFVNRDLAEFSSEL--EYSEMGSSF KSGPKGESAILVEN 306
      || :| :| || | | | : || | | :| :| :| :| :|
Db      7997 PEG---KTQDELDKEAAEAE LNKKVEALPNQVSELEEEELSKLEDNLKDAETNNVEDYIKE 8053

Qy     307 TKEEVIVRSK---DKEDLVCSAALHSPQESVPGKEDRVVSPEKTMDFNEMQMSVVAPVR 363
      || | : :| ||| : : | | | : || | |
Db      8054 GLEEEAIATKQAELEKTPKELDAALN--ELGPDGDEEETPPPE-----APAE 8097

Qy     364 EEEYADFKPFEQAWEVKDTYEGSRDVLAARANVESKVDKCLEL--SLEQKSLGKDSEGRN 421
      : : || |: | | | | | | : | : | | : | :
Db      8098 QPKPE-KPAEETPAPAPKPEKSADQQA-----EEDYARRSEEEYNRLTQQQPPKAEKPAP 8151

Qy     422 EDASFPSTPEPVKDSSRAYITCASFTSA-----TESTTANTFPLLEDHT 465
      | | | | | | | | | | | | | | | |
Db      8152 APAPKPEQPAPA-PKSRGLATKKKLNLAELI LLLKKLGLEPGLEKAGAGLGNLLSTLD 8210

Qy     466 SENKTDE-----KKIEERKAQI--ITEKTSPKTSNPFLVAVQDSEADYVTTDT 511
      | || : | | :| | : : | : | | :| :| :| :|
Db      8211 PEGKTQDELDKEAAEAE LNKKVEALPNQVAELEEEELSKLEDN-----LKDAETNHVEDYI 8265

Qy     512 LSKVTEAAVSNMPE-----GLTPDLVQE---ACESELNEATGTKIAYET--- 552
      : || : | | | | | | : | : | : | | |
Db      8266 KEGLEEEAIATKQAELEKTPKELDAALNELGPDGDEEETPAPEAPAEQPKPEKPAEETPAP 8325

Qy     553 -----KVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSVLPDIVMEAP---- 596
      : | : || | | | | | : | : | : | |
Db      8326 APKPEKSADQQAEEEDYARRSEEEYNRL--TQQQPPKAEKPAPAPAP-KPEQPAPAPKKKQ 8382

Qy     597 -----LNSLLPSA-----GASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAM 641
      | :| | | | : : : | | :| | | :
Db      8383 KVNLENLLSTLDPGGKTQDELDKGAAEAE LNKKVEALPNPVX----ELEEEELSPPEDN-- 8436

Qy     642 NVALKALGT-----KEGIKE-----PESFNAAV-----QETEAPYIS 673
      || | | | :| :| :| :| :| :|
```

```

Db      8437 ---LKDAETNHVEDYIKEGLEEAIATKQAELEETPQEVDAALNDLVDPDGGEEETPAP--- 8490
Qy      674 IACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAE-----LVEDSSPESE 720
      : : : |:| : | || : || | : |:|
Db      8491 ----APQPDEPAPAPAPNAEQPAPAPKPEKSADQQAEEEDYARRSEGEYNRLTQQQPPKAE 8546
Qy      721 -PVDLFSDDSIPEVPQTQEEAVML-----MKESLTVSETVAQHKEERLSASPQ 768
      | : | :| | : : : | | :| :| :|
Db      8547 KPAPAPAPKPEQPAPAPNKEIARLQSDLKDAEENVEDYIKEGLEQAITNKKAEELATTQQ 8606
Qy      769 ELGKPYLESFQPNLHSTKDAASNDIPTLTKE-----KISLQMEEFNIAIYSNDDLL 820
      : | : | | :| | :| : : :| : : :|
Db      8607 NIDKTQKDLEDAELELEKVLATLDPEGKTQDELKAEAEELNEKVEALQNQVAELEEEL 8666
Qy      821 SSKEDKIKESETFSDSSPI-EIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSG 879
      | || :|:| : | :| :| :| :| :|
Db      8667 SKLEDNLKDAETNNVEDYIKEGLEE-----AIATKKAELEKTQKE----- 8706
Qy      880 ADSLPCLELPCDLSEFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIV 939
      | : : | : :| | : | | | :|
Db      8707 -----LDAALNELGPDGD---EEETPAPAPQPEKPAEEPEN-PAPAPKPE----- 8747
Qy      940 KSKSLTKEAEKKLPSDTEKEDRSLS 964
      || :|:| :| :| :|
Db      8748 --KSADQQAEEEDYARRSEEEYNRLT 8770

```

RESULT 9

US-09-976-594-726

; Sequence 726, Application US/09976594

; Patent No. 6673549

; GENERAL INFORMATION:

; APPLICANT: Furness, Michael

; APPLICANT: Buchbinder, Jenny

; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

; FILE REFERENCE: PA-0041 US

; CURRENT APPLICATION NUMBER: US/09/976,594

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/240,409

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 1143

; SOFTWARE: PERL Program

; SEQ ID NO 726

; LENGTH: 2468

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: Incyte ID No. 6673549 4295277CD1

US-09-976-594-726

Query Match 4.9%; Score 288; DB 4; Length 2468;

Best Local Similarity 20.2%; Pred. No. 2e-09;

Matches 251; Conservative 166; Mismatches 441; Indels 386; Gaps 56;

Qy 13 STDSPPR--PPPAFKYQFVTEPEDEEDEEE-----EED-----EEEDDED 50

Db	625	:    :   : :  ::: :   :	684
		ATDVKPKAAKEKTVKKETKVKPEDKKEEKEKPKKEVAKKEDKTPIKKKEEKPKKEEVKKEV	
Qy	51	LEELEVLERKPAAGLSAAAVPPAAAAPLLDfSSDSVPPAPRGPLPAAPPAAPERQPSWER	110
		: ::  :    : ::   : : :::  :	
Db	685	KKEIKKEEKK-----EPKKEVKKETPPKEVKKEVKKEEKKEVKKE---EK	726
Qy	111	SPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRG--	168
		:     :        :     :	
Db	727	EPKKEIKKLPKDAKKSSTPLSEAKKPAALKPKVPKKEESVKKDSVA----AGKPKKEKGKI	782
Qy	169	-----SGSVDETLFAL-----PA-----ASEPVIPISSAEKIMDLME	199
		:       : :   :	
Db	783	KVIKKEGKAAEAVAAVGTGATTAAVMAAAGIAAIGPAKELEAERSLMSSPEDLTkDFEE	842
Qy	200	QPGNTVSSGQEDFPSV-LLETAASLPSLSPLSTVsfKEHGyLGnLSAVSSSEGtIEETLN	258
		: :   :  :      : : :	
Db	843	LKAEEDVDTKDIKPQLELIEDEEKLKETEPEVAYVIQKEREVTKGPAESPDEG-ITTEG	901
Qy	259	EASKELPERATNPFVNRDLAEFSELEYSEMGSsfKGSpgKESAILVENTK-EEVIVRSKD	317
		: : :   :  :  :   :   :     :	
Db	902	EGECEQTPEELEPVEKQGVDDIEKFE--DEGAGFEES--SETGDYEEKAEtEEAEPEED	957
Qy	318	KEDLVC-SAALHSP-----QESPVGKEDRVVSPEKTMDIFNEMQMSV	358
		:      :      :    :      :	
Db	958	GEEHVCVSASKHSPTEDDEESAKAEADAYIREKRESVASGDDRA---EEDMD---EAIEKG	1011
Qy	359	VAPVREEYADFKPFEQAWEVK-DTYEGSRDVLAAANVESKVDRCLEDSLEQ-----	410
		: :  : : :   : :     :    :	
Db	1012	EAEQSEEEADEE--DKAEDAREEEYEPEK--MEAE DYVMAVVDKAAEAGGAEQYgFLTt	1067
Qy	411	--KSLGKDSEGR-----NEDASFPSTPEPVKDSSRAYITCASfTSAtESTT	454
		:   :     :   : :   :	
Db	1068	PTKQLGAQSPGREPASSIHDETLPGGSESEATASDEENREDQPEEFtATSGYTQStIEIS	1127
Qy	455	ANTFPL-----LEDHTSENKTDEK-----KIEERKAQIIITEKTS-PKTSNPFVLVAV	499
		:   : :     : :   : : : : :	
Db	1128	SEPTPMDEMSTPRDVMsDETnNEETESPSQEFVNITKYESSLYSQEYSKPADVTPLNGFS	1187
Qy	500	QDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATGTkIAYETKVDL---	556
		:  :        : :   :   :     :  : :   : :	
Db	1188	EGSKTD--ATDGKYNASASTISPPSSMEED---KFSRSALRDAYCSEVKASTTLdIKDS	1242
Qy	557	--VQTSEAIQESLYPTAQLCP-----SFE-----EAEATPSPVLPDIV	592
		:   :     :   :   :        :    ::	
Db	1243	ISAVSSEKVSpsKSPSLSPSPSPLEKtPLGERSVNFSLTPNEIKVSAEAEVAPVSPEVT	1302
Qy	593	ME-----APLNSLLPSAG-----ASVVQ-----	610
		:   :      : :	
Db	1303	QEVVEEHcASPEDKTLEVVSpsQSVTGSAGHTPYyQSPTDEKSSHLpTEVIEKPPAVPVS	1362
Qy	611	-----PSVSPLEAPPVSYDSIK--LEPENPPP-----YEEAMNVALKALG	649
		: :       :          : :	
Db	1363	FEFSDAKDENERASVSPMDEPVPDESPIEKVLSPLRSPLIGSESAYESfLSADDKASG	1422
Qy	650	TKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPD-FSNYSEIA-----KF	701
		:  :  :     :          :	

Db 1423 -----RGAESPF-----EEKSGKQGSPDQVSPVSEMTSTSLYQDKQ 1458  
 Qy 702 EKSVPPEHAELVEDSSPE-----SEPVDLFSDDSIPEVPQTQ-----EEAV 741  
 | : | : || | | : | : : | | | :  
 Db 1459 EGKSTDFAPIKEDFGQEKKTDDVEAMSSQPALALDERKLGDVSPQTQIDVSQFGSFKEDTK 1518  
 Qy 742 MLMKE-----SLTEVSETVAQ----HKEERLSASPOELG-----KPYLESFQPNLHSTK 786  
 | : | | | | || : | | | : : | : | : ||  
 Db 1519 MSISEGTVSDKSATPVDEGVAEDTYSHMEGVASVSTASVATSSSFPEPTTDDVSPSLH--- 1575  
 Qy 787 DAASNDIPTLTK-KEKISLQMEEFNNTAIYSNDDLSSKEDKIKESETFSDDSSPIEIIDEF 845  
 | | : : : | : : : | : : : ||| : :  
 Db 1576 --AEVGSPHSTEVDDSLSVSVVQTPPT-FQETEMSPSKEECPR----- 1615  
 Qy 846 PTFVSAKDDSPKLAKEYTDLE--VSDKSEIANIQSGADSLPCLELPCDLSFKN----- 896  
 | : | | | | | : : | : | : : | : | | | : :  
 Db 1616 PMSISPPDFSPKTAKSRTPVQDHRSEQSSM-SIEFGQES-PEQSLAMDFSRQSPDHPTVG 1673  
 Qy 897 -----IYPKDE-VHVSDEFSENRSSVSKASISPSNV 926  
 | | : | : | | | : | | |  
 Db 1674 AGVLHITENGPTVDYSPSDMQDSSLSHKIPMEEPSYTDNDLSELISVSQVEASPSTS 1733  
 Qy 927 SALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAE 970  
 || | : : | : : : | : | | | | : : |  
 Db 1734 SAHTP-SQIASPLQEDTLSDVAPPR-----DMSLYASLTSE 1768

RESULT 10

US-09-149-476-411

; Sequence 411, Application US/09149476  
 ; Patent No. 6420526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 186 Human Secreted proteins  
 ; FILE REFERENCE: PZ002P1  
 ; CURRENT APPLICATION NUMBER: US/09/149,476  
 ; CURRENT FILING DATE: 1998-09-08  
 ; EARLIER APPLICATION NUMBER: PCT/US98/04493  
 ; EARLIER FILING DATE: 1998-03-06  
 ; EARLIER APPLICATION NUMBER: 60/040,162  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,333  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/038,621  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,626  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,334  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,336  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/040,163  
 ; EARLIER FILING DATE: 1997-03-07  
 ; EARLIER APPLICATION NUMBER: 60/047,600  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,615  
 ; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,862  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,887  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,908  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/057,650  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/056,884  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,669  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/049,610  
; EARLIER FILING DATE: 1997-06-13  
; EARLIER APPLICATION NUMBER: 60/061,060

; EARLIER FILING DATE: 1997-10-02

Query Match 4.9%; Score 286; DB 4; Length 92;  
Best Local Similarity 56.0%; Pred. No. 2.2e-11;  
Matches 51; Conservative 21; Mismatches 19; Indels 0; Gaps 0;

Qy 1064 NSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIP 1123  
|:|: |:| :| : |||||:||||||| || ||: |||||:||||:||||| : :||:|  
Db 2 NAAMVHINRALKLIIRLFLVEDLVDLSKLAVFMWLMTYVGAVFNGITLLILAELLIFSVP 61  
  
Qy 1124 VIYERHQVQIDHYLGLANKSVKDAMAKIQAK 1154  
::||::: |||||:|:| | : || :|  
Db 62 IVYEKYKTQIDHYVGIARDQTKSIVEKIPSK 92

RESULT 11

US-08-973-462-8

; Sequence 8, Application US/08973462B

; Patent No. 6191270

; GENERAL INFORMATION:

; APPLICANT: DRUILHE, PIERRE

; APPLICANT: DAUBERSIES, PIERRE

; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

; FILE REFERENCE: 0660-0125-0 PCT

; CURRENT APPLICATION NUMBER: US/08/973,462B

; CURRENT FILING DATE: 1998-02-06

; EARLIER APPLICATION NUMBER: PCT/FR96/00894

; EARLIER FILING DATE: 1996-06-12

; EARLIER APPLICATION NUMBER: FR 95/07007

; EARLIER FILING DATE: 1995-06-13

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 1786

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide

US-08-973-462-8

Query Match 4.8%; Score 279.5; DB 3; Length 1786;  
Best Local Similarity 20.1%; Pred. No. 4.4e-09;  
Matches 261; Conservative 220; Mismatches 518; Indels 297; Gaps 60;

Qy 33 EDEEDEEEEEDEEEDDEELEELEVLERKPAAGLSAAAVPPAAAAPLLDFSSDS-----VP 87  
| |: || ||| :|::|| | | | |: :| :| || | |  
Db 225 EKVEESVEENDEESVEENVEE-NVEENDDG---SVASSVEESIASSVDESIDSSIEENVA 280  
  
Qy 88 P-----APRGPLPAAPPAAPERQPSWERSPA-----APAPSLPPAAAV-LPSKLPEDDE 135  
| || || : : || | : | :: :| : : |: |  
Db 281 PTVEEIVAPSVVESVAPSVEESVEENVEESVAENVEESVAENVEESVAENVEESVAENVE 340  
  
Qy 136 PPARP-----PPPPPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALPAASEPVIPS 189  
| | :| : || : : || | : : :| | |  
Db 341 EIVAPTVEEIVAPTVEEIVAPSVVESVAPSVEESVEENVEESVAENV--EESVAENVEES 398  
  
Qy 190 SAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASL--PSLSPLSTVSFKEHGYNLSAVS 247



Db	399	VAENVEESVAE--NVEESVAENVEEIVAPTVEEIVAPTVEEIVAPSVVE-----SVAP	449
Qy	248	SSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGESAILVENT	307
Db	450	SVEESVEENVEESVAENVEESVAENVEESVAENVE-----ES--VAENV	491
Qy	308	KEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYA	367
Db	492	EESV---AENVEESVA----ENVEESVAENVEEIVAP--TVE-----EIVAPTVEEIV	535
Qy	368	DFKPFQEQ-AWEVKDITYEGSRDVLAAARANVESKVDRKCLEDSLEQKSLGKDSEGRN-EDAS	425
Db	536	APSVVESVAPSVEESVEENVEESVAE-NVEESVAEN-VEESVAENV--EESVAENVEEIV	591
Qy	426	FPSTPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTSENKTDEKKIEERKAQIITE	485
Db	592	APTVEEIVAPTVEEIVA----PSVSVESVA----PSVEESVEEN-----VEESVAENVEE	637
Qy	486	KTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQACESELNEATG	545
Db	638	SVAENVEESVAENVEEIVAPTVEEIVAPTVEEIVAPSVVESVAPS-VEESVEENVEESVA	696
Qy	546	TKIAYETKVDLVQTS--EAIQESLYPTAQ--LCPSFEEAEATPSPVLPDIVMEAPLNSLL	601
Db	697	ENVE-ESVAENVEESVAENVEESVAENVEEIVAPTVEEIVA---PTVEEIVAPSVVESVA	752
Qy	602	PSAGASV-----VQPSVSPLEAPPPVSYDSIKLE	630
Db	753	PSVEESVEENVEESVAENVEESVAENVEESVAENVEESVAPTVEEIVAP-----SVE	804
Qy	631	PENPPPYEEAM-----NVALKALGTKEGIKEPESFNAAVQETEAPYISIACDLIK	680
Db	805	ESVAPSVEESVAENVATNLSNLLSLLGGIETEEIKDSILNEIEEVKENVVTTILENVE	864
Qy	681	ETKLSTEPSPDFSNYSEIAK-----FEKSVPEHAELV-----EDSSPESEPVDLF--	725
Db	865	ET--TAESVTTFSNILEEIQENTITNDTIEEKLEELHENVLAALENTQSEEEKKEVIDV	922
Qy	726	-----SDDSIPEVPQTQEEAVMLMKESLTE-----VSETV	755
Db	923	IEEVKEEVATTLIETVEQAEKSANTITEIFENLEENAVESNENVAENLEKLNETVFNTV	982
Qy	756	AQHKEERLSASPQELGKPYLES--FQPNLHSTKDAASNDIPTLT--KEKISLQMEE---	808
Db	983	LDKVEETVEISGESLENNEMDKAFFSEIFDNVKGIQENLLTGMFRSIETSIVIQSEEKVD	1042
Qy	809	FNTAIYSN--DDLSSKEDKIKESETFSDSSPI-EIIDEF-----PTFVS	850
Db	1043	LNENVVSSILDNIENMKEGLLNKLENISSTEGVQETVTEHVEQNVYVDVDVPAMKDQFLG	1102
Qy	851	AKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPKDEVHVSDEFS	910
Db	1103	ILNEAGGLKEMFFNLEDVFKSE-----SDVITVEEIKDEPVQKEV-EKETVSIIEEME	1154
Qy	911	ENRSSV---SKASISPSNVSALEPQTEMGSIKSKSLT-KEAEKKLPDTEK-----EDR	961

Db 1155 ENIVDVLEEEKEDLTDKMIDAVEESIEISSDSKEETESIKDKKDVSLVVEEVQDNDMDE 1214  
 Qy 962 SLSAVL-----SAELSKTSV-VDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIA 1015  
 | : | | | : | : : : | : | : | :  
 Db 1215 SVEKVLELKNMEEELMKDAVEINDITSKLIEETQELNEVEADLIKDME-----K 1263  
 Qy 1016 LALLSVTISFRIYKGVQA----IQK-SDEGHPFRAYLESEVA---ISEELVQKYSN---S 1065  
 | | : | | : | | : | | : | | : | : | :  
 Db 1264 LKELEKALS-EDSKEIIDAKDDTLEKVVIEEHDTTLDDEVVELKDVEEDKIEKVSDLKD 1322  
 Qy 1066 ALGHVNSTIKELRRLFLVDDLVDLSLKFVAVLMMVFTYVGALFNGLTLLILALISLFSIPVI 1125  
 : : | : | : : : : : : : : :  
 Db 1323 LEEDILKEVKEIKE--LESEILEDYK-----ELKTIETDIL 1356  
 Qy 1126 YERHQVQIDHYLGLANKS--VKDAMAKIQAKIPGLK 1159  
 | : : : | : : : | : | : : | :  
 Db 1357 EEKKEIEKDHFEKFEEEAEEIKDLEADILKEVSSLE 1392

RESULT 12

US-08-769-309A-5

; Sequence 5, Application US/08769309A

; Patent No. 5741890

; GENERAL INFORMATION:

; APPLICANT: Scott, John D.,

; APPLICANT: Nauert, Brian J.,

; APPLICANT: Klauck, Theresa M.

; TITLE OF INVENTION: Protein Binding Domains of Gravin

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower/233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/769,309A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5741890and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27866/33451

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1780 amino acids

; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-769-309A-5

Query Match 4.5%; Score 265.5; DB 1; Length 1780;  
Best Local Similarity 20.8%; Pred. No. 3.4e-08;  
Matches 264; Conservative 146; Mismatches 486; Indels 371; Gaps 51;

```
Qy      11 SSSTDSPPRPPPA-----FKYQFVTE-----PEDE----EDEEEEEDEEED 47
      || :|| | : || :| : ||| | :|| :|
Db      277 SKSAESPTSPVTSETGSTFK-KFFTQGWAGWRKKTSTRKPKEDVEASEKKKEQEPEKVD 335

Qy      48 DEDLEEEVLERKPAAGLSAAAVPPAAAA--PLLDFSSDSVPPAPRGPLPAAPPAAPERQ 105
      |: : || | | | || :| | | : | ||: | |
Db      336 TEEDGKAEVASEKLTASEQAHPQEPAESAEHPRLSAEYEKV-----ELPS-----EEQ 383

Qy     106 PSWERSPAAPAPSLPPAAAVLPSKLP-----EDDEPPARPPPPPPAGASP 150
      | : |: |: | | | |: | : : ||: |
Db      384 VSGSQGPSEEKPA-PLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQKTEVEETAGSVP 442

Qy     151 L-----AEP--AAPPSTPAAPKRRSGSVDETLFALPAASEPVI PSSAEKIMDLMEQP 201
      ||| | | | | | | : | |: | :: :|
Db      443 AEELVGMDAEPQEAEPAKELVKLKETCVSGEDPTQGADLSPDEKVL SKPPEGVVSEVEML 502

Qy     202 GNTVSSSQEDFP SVLLETAASLPSLSPLSTVSFKEHGYLG-----NLSAVSSSEGTI 253
      : : | | |: | || |: | | : : | :
Db      503 SSQERMKVQGSPLKKLFTSTGLKKLS-----GKKQKGRGGGDEESGEHTQVPADSPDSQ 557

Qy     254 EETLNEASKELPERATN-PFVNRDLAEFSELEYSEMGS SFKGPSKGESAILVENTKEEVI 312
      || |:| || : : ||| : :| |: | | | : |: |
Db      558 EEQKGESSASSPEEPEEITCLEKGLAEVQQDGEAE EGATSDGEKKREGVTPWASFKKMVT 617

Qy     313 VRSK-----DKE---DLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVAPV 362
      : : ||| | | || | : : : : ||: || |
Db      618 PKKRVRRPSESDKEDEL DKVKSATLSSTEST-----ASEMQEEMKGSVEEPK 664

Qy     363 REEYADFKPFEQAWEV-----KDTYEGS 385
      || :|| ||| | | | | | | | | :
Db      665 PEEPKRKVDTSVSWEALICVGSSKKRARRRRSSSDEEGGPKAMGGDHQKADEAGKDKETGT 724

Qy     386 RDVLA-----ARANVESK-----VDRKCLED SLEQKSLGKDS 417
      :|| | : | : || : ||: || :||
Db      725 DGILAGSQEHDPGQGSSSPEQAGSPTEGEGVSTWESFKRLVTPRKKS SKSLEEK--EDS 782

Qy     418 -EGRNEDASFPSTPEPVKDSS-----RAYITCASFTSATE--ST 453
      | : | | | | |: | :| : | | | |
Db      783 IAGSGVEHSTPDT-EPGKEESWVSIKKFIPGRRKKRPD GKQEAPVEDAGPTGANEDDSD 841

Qy     454 TANTFPILLEDHTSENKTDE-----KKIEERKAQIITEKTS PKTSNPFVLVAVQD---- 501
      || | | : | : : |:| : :|| : || |
Db      842 VPAVVPLSEYDAVEREKMEAQQAQKGAEQPEQKAATEVSKELSESQVHMAAAVADGTRA 901

Qy     502 -----SEADYVTTDTLSK-----VTEAAVSNMPEG- 526
      :|| :| :| : ||| :||
Db      902 ATII EERSPSWISASVTEPLEQVEAEAALLTEEVLEREVIAEEEPPTVTEP----LPENR 957

Qy     527 -----LTPDLVQEA-CESELNEATGT KI-AYETKVDLVQTSEAIQESLYPTAQ 572
```

Db	958	EARGDTVVSEAEALTPEAVTAAETAGPLGSEEGTEASAAEETTEMVSAVSQLTDSPTDTEE	1017
Qy	573	L-----CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVS	623
Db	1018	ATPVQEVVEGGVPDIEEQERRTQEVQLQAVAEKVKEESQLPGTG-----GPEDVLQPVQ	1069
Qy	624	YDSIKLEPENPPPYEEA-----MNVALK-----ALGTKEGIKEPESFNAAVQE	666
Db	1070	-----RAEAERPEEQAEASGLKKETDVVLKVDAQEAKTEPFTQGKVVGGQTTPESEKAPQV	1125
Qy	667	TEAPYISIACDLIKETKLSTEPSPD--FSNYSEIAKFEKSVPEHAELVEDSSPES--EPV	722
Db	1126	TES-----IESSELVTTCQAEATLAGVKSQEMVMEQAIP-----PDSVETPT	1166
Qy	723	DLFSDDSSIP----EVPQTQEEAVMLMKESLTVSETVAQH----KEERLSASPQELGKPY	774
Db	1167	DSETDGSTPVADFDAPGTTQ-----KDEIVEIHEENEVHLVPVRGTEAEAVPAQKERPP	1220
Qy	775	LESFQPNLHSTKDAASNDIPTLTCKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETF	834
Db	1221	APSSFVFQEETKE-----QSKMEDTLEHTDKEVSVETVSILSKTEGTQEADQYA	1269
Qy	835	DSSPIEI-----IDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPC	885
Db	1270	DEKTKDVPFFEGLEGSIDTGITVSREKVTEVALKGGEGTEEAECKKDDALELQSHAKSPPS	1329
Qy	886	LELPCDLSF-----KNIYPKDEVHVSDEFSENRSS-----VSKASISPSNVSALEPQTE	934
Db	1330	---PVEREMVQVEREKTEAEPHTVNEEKLEHETAVTVSEEVSKQLLQTVNVPIDGAKE	1386
Qy	935	MGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVF--	992
Db	1387	VSSLEGSPPPCLGQEEAVCTKIQVQSSEASFTLTAAAEEEKV--LGETANILETGETLEP	1444
Qy	993	-GASLFL	998
Db	1445	AGAHVL	1451

RESULT 13

US-08-994-570-5

; Sequence 5, Application US/08994570

; Patent No. 6090929

; GENERAL INFORMATION:

; APPLICANT: Scott, John D.,

; APPLICANT: Nauert, Brian J.,

; APPLICANT: Klauck, Theresa M.

; TITLE OF INVENTION: Protein Binding Domains of Gravin

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower/233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-994-570-5

```

```

Query Match          4.5%; Score 265.5; DB 3; Length 1780;
Best Local Similarity 20.8%; Pred. No. 3.4e-08;
Matches 264; Conservative 146; Mismatches 486; Indels 371; Gaps 51;

```

```

Qy      11 SSSTDSPPRPPPA-----FKYQFVTE-----PEDE----EDEEEEEDEEED 47
      ||:|| | : ||:| | : ||| | :||| | :|
Db      277 SKSAESPTSPVTSETGSTFK-KFFTQGWAGWRKKTsFRKPKEDEVEASEKKKEQEPEKVD 335

Qy      48 DEDLEELEVLERKPAAGLSAAAVPPAAAA--PLLDfSSDSVPPAPRGPLPAAPPAAPERQ 105
      |: :|| | | | ||:| | | : | ||: | |
Db      336 TEEDGKAEVASEKLTASEQAHPOEPAESAHEPRLSAEYEKV-----ELPS-----EEQ 383

Qy      106 PSWERSPAAPAPSLPPAAAVLPSKLP-----EDDEPPARPPPPPPAGASP 150
      | : |: |: | | | | : | : | | |: |
Db      384 VSGSQGPSEEKPA-PLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQKTEVEETAGSVP 442

Qy      151 L-----AEP--AAPPSTPAAPKRRGSGSVDETlFALPAASEPVIPSSAEKIMDLMEQP 201
      ||| || | | | | | | | : | | : | :|
Db      443 AEELVGMDAEPQEAEPAKELVKLKETCVSGEDPTQGADLSPDEKVLSPPEGVVSEVEML 502

Qy      202 GNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYLG-----NLSAVSSSEGTI 253
      : : | | |: | || | : | | : : | :
Db      503 SSQERMKVQGSPLKKLFTSTGLKKLS-----GKKQKGKRGGGDEESGEHTQVPADSPDSQ 557

Qy      254 EETLNEASKELPERATN-PFVNRDLAEFSELEYSEMGSsfKGSPKGESAILVENTKKEEVI 312
      || |:| || : : ||| : :| |: | | | : |: |
Db      558 EEQKGESSASSPEEPPEEITCLEKGLAEVQQDGEAEEGATSDGEKKREGVTPWASFKKMT 617

Qy      313 VRSK-----DKE---DLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVPV 362
      : : ||| | | | | | : : : : |: | | |
Db      618 PKKRVRRPSESDKEDELdkVKSATLSSTEST-----ASEMQEEMKGSVEEPK 664

```

Qy 363 REEYADFKPFQAWEV-----KDTYEGS 385  
 || :|| || :|  
 Db 665 PEEPKRKVDTSVSWEALICVGSSKKRARRRSSSDEEGGPKAMGGDHQKADEAGKDKETGT 724

Qy 386 RDVLA-----ARANVESK-----VDRKCLEDSEQKSLGKDS 417  
 :|| | : | : || : ||:| :||  
 Db 725 DGILAGSQEHDPGQGSSSPEQAGSPTEGEGVSTWESFKRLVTPRKKSkskLEKS--EDS 782

Qy 418 -EGRNEDASFPSTPEPVKDSS-----RAYITCASFTSATE--ST 453  
 | : | | | | : | :| : | | | |  
 Db 783 IAGSGVEHSTPDT-EPGKEESWVSIKKFIPGRRKKRPDGKQEQAPVEDAGPTGANEDDSD 841

Qy 454 TANTFPLLEDHTSENKTDE-----KKIEERKAQIITEKTSPKTSNPFLVAVQD----- 501  
 || | | : | :| :| :| :| :| :|  
 Db 842 VPAVVPLSEYDAVEREKMEAQQAKGAEQPEQKAATEVSKELSESQVHMAAAVADGTRA 901

Qy 502 -----SEADYVTTDTLSK-----VTEAAVSNMPEG- 526  
 :|| :| :| :| :| :| :|| :||  
 Db 902 ATIIERSPSWISASVTEPLEQVEAEALLTEEVLEREVIAEEEPPTVTEP-----LPENR 957

Qy 527 -----LTPDLVQEA-CESELNEATGTKI-AYETKVDLVQTSEAIQESLYPTAQ 572  
 |||: | | | :| :| :| :| :| :| :|  
 Db 958 EARGDTVSEAEALTPEAVTAAETAGPLGSEEGTEASAAEETTEMVSAVSQLTDSPTDTEE 1017

Qy 573 L-----CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVS 623  
 | || | | : : | || | | : ||  
 Db 1018 ATPVQVEGGVPDIEEQERRTQEVLLQAVAQKVKESQLPGTG-----GPEDVLQPVQ 1069

Qy 624 YDSIKLEPENPPPYEEA-----MNVALK-----ALGKKEGIKEPESFNAAVQE 666  
 : | | | | :| :| | | | | | | |  
 Db 1070 ----RAEAERPEEQAEASGLKKETDVVLKVDQAQEAKEPFTQGKVVGGQTTPESEKAPQV 1125

Qy 667 TEAPYISIACDLIKETKLSTEPSPD--FSNYSEIAKFEKSVPEHAELVEDSSPES--EPV 722  
 ||: | : :| | : | :| :| :| :| :|  
 Db 1126 TES-----IESSELVTTCAETLAGVKSQEMVMEQAIP-----PDSVETPT 1166

Qy 723 DLFSDDSSIP----EVPQTQEEAVMLMKESLTVSETVAQH----KEERLSASPQELGKPY 774  
 | :| | | : | | : | :| :| | : | :| :|  
 Db 1167 DSETDGSTPVADFDAPGTTQ-----KDEIVEIHEENEVHLVPVRGTAEAVPAQKERPP 1220

Qy 775 LESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETF 834  
 | ||: | : | :| :| :| :| :| :| :| :| :|  
 Db 1221 APSSFVFQEETKE-----QSKMEDTLEHTDKEVSVETVSILSKTEGTQEQADQYA 1269

Qy 835 DSSPIEI-----IDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPC 885  
 | : : || | | | | : | : || | | |  
 Db 1270 DEKTKDVPFFEGLEGSIDTGITVSREKVTEVALKGEGTEEAECKKDDALELQSHAKSPPS 1329

Qy 886 LELPCDLSF-----KNIYPKDEVHVSDEFSENRSS-----VSKASISPSNVSALEPQTE 934  
 | : : : ||:| | : : ||| : || : : |  
 Db 1330 ---PVEREMVVQVEREKTEAEPHTVNEEKLEHETAVTVSEEVSKQLLQTVNVPIDGAKE 1386

Qy 935 MGSIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVF-- 992  
 : | : | : : : : | | : | | :| :||  
 Db 1387 VSSLEGSPPPCLGQEEAVCTKIQVQSSEASFTLTAAAEKVK--LGETANILETGETLEP 1444

Qy 993 -GASLFL 998

|| | |  
Db 1445 AGAHLVL 1451

RESULT 14

US-09-134-001C-4463

; Sequence 4463, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4463

; LENGTH: 2137

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4463

Query Match 4.5%; Score 265; DB 4; Length 2137;

Best Local Similarity 19.6%; Pred. No. 4.8e-08;

Matches 172; Conservative 141; Mismatches 451; Indels 114; Gaps 24;

```
Qy      159 STPAAPKRRGSGSVDETLFALPAASEFPVIPSSAEKIMDLMEQPNTVSSGQEDFPSVLLE 218
      || :   | | |   : | :   | : |   :   : | : |   :   |
Db      990 STSTSTSDSASTSTSE-----SES DSASTSLSESTSTSVSDSTSTSTSDSASMSASESE 1043

Qy      219 TAASLPSLSPLSTVSFKEHGYLG NLSAV---SSSEGTIEETLNEASKELPERATNPFVNR 275
      : :   ||| : : |   | : ||   | : |   | |   :   | | | :
Db     1044 SNSKSTSLSESTSTSL S-----GSTSASTSDSASTSTSESESDSTSTSLSESTSTSLSGS 1098

Qy      276 DLAEFSELEYSEMGS SFGKSPKGESAILVE-----NTKEEVIVRSKDKEDLVC 323
      | : |   | : :   | | |   | | |   : | |   : : |
Db     1099 TASTSD---SASTSTSESDSTSESTSLSESLSTSVSDSTSASTSESASTSTSESESN SA 1155

Qy      324 SAALHSPQESPVGKEDRVVSPEKTM DIFNEMQMSVVAPVREEYADFKPFQAWEVKDTYE 383
      | : |   : :   : :   : | :   :   |   : : |
Db     1156 STSLSGSLSTSI SDSTSTSTSDSASTSTSESESDSTSTSLSE-----STSTSLSDSTS 1208

Qy      384 GSRDVLAA RANVESKVDRKCLED SLEQK---SLGKDSEGRNEDASFPSTPEPVKDSSRAY 440
      |   | : :   | | |   | | :   | :   :   | : |   : :
Db     1209 TSTSESASTSTSES--DSTSESTSLSESTSTSVSDSTSASTSDSASTSTSVSDSESA STS 1266

Qy      441 ITCASFTSATESTTANTFPLLEDHTSE--NKTDEKKIEERKAQIITEKTS PKTSNPFLVA 498
      | : :   | | : | : : |   | | |   : : :   : | :   : : | | :
Db     1267 ISESLSTSVSDSTSTSTSDSASTSTSESDSTSESTSLSE SISTSVSDSTSASTSDSASTS 1326

Qy      499 VQDSEADYVTTD-----T LSKVTEAAVSNMPEGLTPDLVQEACESELNEATG TKIAYE 551
      : | : |   : |   : | |   | : |   | :   | : : | | : :
```

Db 1327 TSESESDSASTSLSGSTSTSLSDSTSTSTSDSASTSTSESDSERASTSLSGSTSTSLSDS 1386  
 Qy 552 TKV---DLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASV 608  
 | | | | : : | | : | | : : | | |  
 Db 1387 TSTSTSDSASTSTSVSDS-----NSASTSLSGSLSTSVSDSTSTSTSDSASAST 1435  
 Qy 609 VQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALK-ALGTKEGIKEPESFNAAVQET 667  
 | | | | : : : : : | : | | : : | :  
 Db 1436 ---SESDSERA-----STSLSGSTSTSTSDSTSTSTSDSASTSTSVSESNSTSTSTSES 1486  
 Qy 668 EAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSD 727  
 : : | | | | : : : : : | | | : : | :  
 Db 1487 LSTSVS-----DSTSTSTSDSASTSTSVSDSDSASTSSSESV--STSDSESTSTST 1536  
 Qy 728 DSI---PEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQE-----LGKPYLESFQ 779  
 | | | : : | : | | : : : : | | | |  
 Db 1537 DSASTSTSVSESNSTSTSLSGSTSTSVSDSTSTSTSDSASASTSESDSDSASTSSSESVS 1596  
 Qy 780 PNLH-----STKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFS 834  
 : : | | : : | : | : : : : : : : | : |  
 Db 1597 TSVSDSTASTSEASTSTSVSDSNSASTSLSESTSTSLSDSTSMSTSDSASTSTSESDS 1656  
 Qy 835 DSSPIEIID-----EFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADS 882  
 | : : | | | : : : : : | : : : : | :  
 Db 1657 DSASTSLSDSTSTSVSESTSTSTSTSVSASNSTSTSLSDSRSTSLSDSTSTSTSESGSTS 1716  
 Qy 883 LPCLELPCDLSEFKNIYPKDEVHVSDEFSENRR--SSVSKASISPSNVSALEPQTEMGSIVK 940  
 | | : : : | | | : : | : : | : : : : | :  
 Db 1717 TS--ESDSDSASTSLSESTSTSTSDSTSTSTSDSASTSMSVSDSNRASTSLSDSTSTSVS 1774  
 Qy 941 SKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVD 978  
 : : | | | | | | : | | |  
 Db 1775 DSTASTSEASTSTRESESTASTSLSESTSTSVSD 1811

RESULT 15

US-09-621-976-4600

; Sequence 4600, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 4600

; LENGTH: 75

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 58

; OTHER INFORMATION: Xaa = His,Pro



; NAME/KEY: UNSURE  
; LOCATION: 28  
; OTHER INFORMATION: Xaa = Met,Val  
; NAME/KEY: UNSURE  
; LOCATION: 19  
; OTHER INFORMATION: Xaa = Pro,Gln  
; NAME/KEY: UNSURE  
; LOCATION: 53  
; OTHER INFORMATION: Xaa = Ser,Tyr  
US-09-621-976-4600

Query Match 4.4%; Score 256.5; DB 4; Length 75;  
Best Local Similarity 67.0%; Pred. No. 1.2e-09;  
Matches 59; Conservative 3; Mismatches 11; Indels 15; Gaps 3;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
          |||:||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MEDLDQSPLVSSS-DSPPRXQPAFKYQFXREPEDEE-----EDLEEELEVLERK 47

Qy      61 PAAGLSAAAV--PPAAAAPLLDFSSDSV 86
          ||||| || | ||| |||:||| :| |
Db      48 PAAGLXAAPVXTAPAAGAPLMDFGNDFV 75
```

Search completed: September 29, 2004, 18:20:58  
Job time : 37.8717 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:05:43 ; Search time 28.465 Seconds  
(without alignments)  
3930.111 Million cell updates/sec

Title: US-09-830-972-2  
Perfect score: 5848  
Sequence: 1 MEDIDQSSSLVSSSTDSPPRP.....VKDAMAKIQAKIPGLKRRKAD 1163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	789.5	13.5	776	2	A46583	neuroendocrine-spe
2	685	11.7	208	2	I60904	neuroendocrine-spe
3	671	11.5	267	2	A60021	tropomyosin-relate
4	517	8.8	2484	2	T26216	hypothetical prote
5	503.5	8.6	2607	2	T26215	hypothetical prote
6	328.5	5.6	5327	2	T13564	microtubule-associ
7	322	5.5	7962	2	I38346	elastic titin - hu
8	320	5.5	222	2	T26213	hypothetical prote
9	304.5	5.2	873	2	A47283	calphotin - fruit
10	302.5	5.2	1829	2	T24583	hypothetical prote
11	295.5	5.1	865	2	A47282	calcium-binding pr
12	292	5.0	3507	2	T34513	hypothetical prote
13	291.5	5.0	2364	2	A56577	microtubule-associ

14	288.5	4.9	971	2	T19431	hypothetical prote
15	284	4.9	2464	1	QRMSPI	microtubule-associ
16	281	4.8	3488	2	T34418	hypothetical prote
17	279.5	4.8	1262	2	T22523	hypothetical prote
18	277	4.7	1621	2	A82255	hypothetical prote
19	275.5	4.7	3924	2	S37431	ankyrin 2, neurona
20	275	4.7	1299	2	T47182	hypothetical prote
21	273.5	4.7	1029	2	T30351	mucin-like protein
22	273.5	4.7	1274	2	T16251	hypothetical prote
23	273	4.7	1558	2	B71603	RESA-H3 antigen PF
24	272.5	4.7	3534	2	T42567	tegument protein 2
25	270	4.6	1230	2	T22458	hypothetical prote
26	269.5	4.6	2187	2	T30826	nascent polypeptid
27	267	4.6	1684	2	JW0057	gravin - human
28	265.5	4.5	1828	2	A40115	microtubule-associ
29	263	4.5	1825	2	S13507	microtubule-associ
30	261.5	4.5	1087	1	QFMSH	neurofilament trip
31	261	4.5	606	2	A43427	neurofilament trip
32	261	4.5	2570	2	T17451	fimbriae-associate
33	260	4.4	1804	2	T34518	nestin - golden ha
34	259.5	4.4	1020	1	QFHUH	neurofilament trip
35	259	4.4	1510	2	T33100	hypothetical prote
36	258.5	4.4	1830	2	A37981	microtubule-associ
37	257	4.4	1224	2	T14007	microtubule-associ
38	256	4.4	2361	2	T25752	hypothetical prote
39	254.5	4.4	6642	2	T29757	protein UNC-89 - C
40	254	4.3	1616	2	G64242	cytadherence-acces
41	254	4.3	3381	2	T42389	versican precursor
42	253	4.3	3421	1	WZBEB6	367K tegument prot
43	252.5	4.3	5170	2	T15348	hypothetical prote
44	251	4.3	1824	1	QRHUMT	microtubule-associ
45	250.5	4.3	4377	2	A55575	ankyrin 3, long sp

# ALIGNMENTS

## RESULT 1

A46583

neuroendocrine-specific protein, splice form A - human

N;Contains: neuroendocrine-specific protein, splice form B

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 05-Nov-1999

C;Accession: A46583; I60903

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: A46583

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-776 <ROE1>

A;Cross-references: GB:L10333; NID:g307306; PIDN:AAA59950.1; PID:g307307

A;Accession: I60903

A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 421-776 <ROE2>  
A;Cross-references: GB:L10334; NID:g307308; PIDN:AAA59951.1; PID:g307309  
C;Genetics:  
A;Gene: GDB:RTN1; NSP  
A;Cross-references: GDB:203968; OMIM:600865  
A;Map position: 14q21-14q22

Query Match 13.5%; Score 789.5; DB 2; Length 776;  
Best Local Similarity 31.2%; Pred. No. 3.5e-25;  
Matches 243; Conservative 115; Mismatches 252; Indels 169; Gaps 28;

```

Qy      487 TSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATGT 546
      : | : | ||: : : : | : | || : : | :
Db      65 SGPARQSP--VAMETASTGVAGVSSAMDHTFSTTSKDGEG-----SCYTSLI----S 110

Qy      547 KIAETKVDLVQTSEAIQ-ESLYPTAQLCPSEEEAEATPSPVLPDI--VMEAPLNS---- 599
      | | : | : : | : | | | ||: : | |
Db      111 DICYPQEDSTYFTGILQKENGHVITISESP---EELGTPGPSLPDVPGLIESRGLFSSDSG 167

Qy      600 --LLPSAGASVVQPSVSPLEAPPPVSY-----DSIKLEPENPPPPYEEA-----M 641
      : | : | : ||: : | : : | : : | :
Db      168 IEMTPAESTEVNKILADPLDQMAEAYKYIDITRPEEVKHQEQHHPELEDKDLDFKNKDT 227

Qy      642 NVALKALGTKEGIKEPE-----SFNAAVQETEAPYISIACDLIKETKLSTE-PSP 690
      : : : | ||: ||: : : : | ||| || : | : : : |
Db      228 DISIK----PEGVREPDKPAPVEGKIIKDHLLEESTFAPYID---DLSEEQRRAPQITTP 280

Qy      691 DFSNYSEIAKFEKSVPEHAELVEDSSPESEPVDLF-SDDSIPEV----PQTQEEAVMLMK 745
      : || : : : ||: : : | | : | | : :
Db      281 VKITLTEIE-----PSVETTTQEKTPKQDICKPSPDTVPTVTVSEPEDDSPGSITPP 334

Qy      746 ESLTEVSETVAQHK-----EERLSASPQELGKPYLESFQP---NLHSTKDAASND---- 792
      | || | : | : : | : | : | : : :
Db      335 SSGTEPSAAESQKGKSISEDELITAIKEAKGLSYETAENPRPVGQLADRPEVKARSGPPT 394

Qy      793 IPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKI----- 827
      ||: | | : : : | | : : || :
Db      395 IPSPLDHEASSAESGDSEIELVSEDPM--AAEDALPSGYVSFGHVGGPPPPSPASPSIQYS 452

Qy      828 ---KESETFSDSSPIEIIDFPTFVSAKDDSPKLAKEYTDLEVSDKSEI-----ANIQSG 879
      : | | || | : | | || : ||| : : : | | :
Db      453 ILREEREAE LDS---ELIIESCDASSASEESPKREQDSPPMKPSALDAIREETGVRAEER 509

Qy      880 ADSLPCLELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIV 939
      | | | | | : || | | | | | | | | | |
Db      510 APSRRGLAEPG--SFLD-YPSTEPQPGPEL-----PPGDGALEPETPM---- 549

Qy      940 KSKSLTKEAEKKLPDTEKEDRSLSAVLSAE-----LSKTSVVDLLYWRD 984
      || | : : | : : : | : : | | | | |
Db      550 -----LPRKPEEDSSSNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLLYWRD 597

Qy      985 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPF 1044
      ||: ||: ||: | | || | ||: || | || | | | | | | | | | |
Db      598 IKQTGIVFGSFLLLLFSLTQFSVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPF 657

```

```
QY      1045 RAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMMWVFITYVGA 1104  
       :||| |: :||: |||::   :|||:||||| | ||||| |||||: ||||  
Db      658 KAYLELEITLSQEIQKYTDCLQFYVNSTLKELRRLFLVQDLVDLSLKFAVLMMWLLTYVGA 717  
  
QY     1105 LFNGLTLLILALISLFSIPVIYERHQVQIDHYLGANKSVKDAMAKIQAKIPGLKRKAD 1163  
       |||||||::|::|::|::|: || ||| |||   :   :||| |||| || |:  
Db     718 LFNGLTLLLMAVVSMTFLPVVYVKHQAQIDQYLGLVRTHINAVVAKIOAKIPGAKRHAE 776
```

## RESULT 2

I60904

neuroendocrine-specific protein C - human

C;Species: Homo sapiens (man)

C;Date: 24-May-1996 #sequence revision 24-May-1996 #text change 05-Nov-1999

C;Accession: I60904

R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A;Reference number: A46583; MUID:93293865; PMID:7685762

A;Accession: I60904

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-208 <RES>

A;Cross-references: GB:L10335; NID:g307310; PIDN:AAA59952.1; PID:g307311

## C; Genetics:

A;Gene: GDB:RTN1; NSP

A;Cross-references: GDB:203968; OMIM:600865

A;Map position: 14q21-14q22

Query Match 11.7%; Score 685; DB 2; Length 208;

Best Local Similarity 67.0%; Pred. No. 1e-21;

Matches 128; Conservative 32; Mismatches 31; Indels 0; Gaps 0;

Qy 973 KTSVVDLLYWDRDIKKTGVVFGLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGI 1032  
| : : ||||| | : || : | || || | : || | | : || | | : ||||| | :  
Db 18 KSQAIDLLYWDRDIQTGIVFGSLFLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 77

Qy 1033 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKF 1092  
||:|:|||||:|||| |: :||: |:||:| :|||:|||||:||||| |||||  
Db 78 QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQFYVNSTLKLRLFLVQDLVDLSLKF 137

Qy 1093 AVLMWVFYTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ 1152  
 ||||: |||||::||::||::||: || ||| |||| : |||  
 Db 138 AVLMWLLTYVGALFNGLTLLMAVVSMTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQ 197

```
QY      1153 AKIPGLKRKAD 1163
          ||||| || |:
```

Db 198 AKIPGAKRHAE 208

### RESULT 3

A60021

tropomyosin-related protein, neuronal - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 05-Nov-1999  
 C;Accession: A60021  
 R;Wieczorek, D.F.; Hughes, S.R.  
 Brain Res. Mol. Brain Res. 10, 33-41, 1991  
 A;Title: Developmentally regulated cDNA expressed exclusively in neural tissue.  
 A;Reference number: A60021; MUID:91278684; PMID:1647480  
 A;Accession: A60021  
 A;Molecule type: mRNA  
 A;Residues: 1-267 <WIE>  
 A;Cross-references: EMBL:X52817; NID:g456549; PIDN:CAA37001.1; PID:g456550  
 C;Comment: This neuronal-specific mRNA was identified by hybridization to an alpha-tropomyosin probe but does not show homology in amino acid sequence.

Query Match 11.5%; Score 671; DB 2; Length 267;  
 Best Local Similarity 66.3%; Pred. No. 5.6e-21;  
 Matches 124; Conservative 33; Mismatches 30; Indels 0; Gaps 0;

```

Qy      973 KTSVVDLLYWDRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKQVI 1032
          | : :| | | | | | | | :| | :| | : | | | | | | | | | | | | | | :
Db      9   KSQAIDLLYWDRDIKQTGIVFGSFLLLLSLTQFSVSVVAYLALALSAISFRIYKSVL 68

Qy     1033 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLK 1092
          | | :| | :| | | | | :| | :| | :| | :| | :| | :| | | | | | | |
Db      69   QAVQKTDEGHPFKAYLELEITLSQEIQKYTDCLQLYVNSTLKELRRLFLVQDLVDLSLK 128

Qy     1093 AVLMMWVFETYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQ 1152
          | | | | : | | | | | | | | | :| | :| | :| | :| | :| | :| | | | |
Db     129   AVLMWLLTYVGALFNGLTLLLMAVSMFTLPVVVYVKHQAQVDQYLGLVRTHINTVVAKIQ 188

Qy     1153 AKIPGLK 1159
          | | | | :
Db     189   AKIPGAR 195
  
```

#### RESULT 4

T26216

hypothetical protein W06A7.3c - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T26216

R;Ainscough, R.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20173

A;Accession: T26216

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-2484 <WIL>

A;Cross-references: EMBL:Z78066; PIDN:CAB51467.1; GSPDB:GN00023; CESP:W06A7.3c

A;Experimental source: clone W06A7

C;Genetics:

A;Gene: CESP:W06A7.3c

A;Map position: 5

A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2

Query Match 8.8%; Score 517; DB 2; Length 2484;  
 Best Local Similarity 20.1%; Pred. No. 2.7e-13;  
 Matches 262; Conservative 189; Mismatches 413; Indels 440; Gaps 51;

Qy 28 FVTEPEDEE---DEEEEEDEEED-----DEDLEELEVL----- 57  
 | | | : | : | | | | | | : | : |  
 Db 1437 FGTSSEESQKADGNQENQEEEDVVAELNFHPIRQWRDEDVISLQSLKSLVAEVCITDV 1496  
 Qy 58 -----ERKPAAGLSAAAVPPAAAAPL-LDFSSDSVPPAPRGPLPAAPPAAP--ERQPSW 108  
 | : : | : | : | | : | | : | | |  
 Db 1497 DASDVNEQDEESTLKILKVVPSEPSLLELDFTND--PKVIHVPIPLMEPATMYLEEMVEW 1554  
 Qy 109 -----ERSPAAPAPSLPPAAAVLPS-----KLP-EDDEPPARPPPPPPA 146  
 | : | | : : : | | | | | | | |  
 Db 1555 IIADAVKEVSEMEVVTSEISEMAPQVSESTCPIPEPLADLKLPVEDDEKTPEPEPVVPG 1614  
 Qy 147 GASPLAEPAPPSTPAAPKRRGSGSVDETFLALPAASEPVIPS--SAEKIMDLMEQPGNTV 205  
 | | | | : | | | : | : | : | : |  
 Db 1615 QVQERIPIEVEQAPTIPQRPPRAPKSE----LPKVAKPLDDSKSRVRFAPLNIKLGRTY 1670  
 Qy 206 SSGQEDFPSVLLETAASLPSLSPLSTVSKHEGYLGNLSAVSSSEGTEETLNEASKELP 265  
 | | : | | | | : : : : :  
 Db 1671 SEEQQK-----ELVESLE--RPLTIIT-----QQKPP 1695  
 Qy 266 ERAT-----NPFVNRDLAEFSELEYSEMGSFSGSPKGESAILVENTKEEVIVRSKDKE 319  
 | : | : | | | : : | : |  
 Db 1696 EKPTEDIGALSPLSPNTLAEYEEVPMMDMQS----- 1726  
 Qy 320 DLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQ-MSVVAPVREEYADFKPFQAWEV 378  
 | | | | | : : | : | | | | | | |  
 Db 1727 -----VPHSPQE----KQEEI---EALSEIIEEPQAMKEVE-----KPVESAPE- 1763  
 Qy 379 KDTYEGSRDVLAAARANVESKVDRKCLEDL--EQKSLGKDSEGRNEDASFPSTPEPVKDS 436  
 | | : | | : : | : | : | | : | : |  
 Db 1764 KD-----NESLEAPEIINEPIRRVLVETKIMGPGKSLNED---NDD----- 1801  
 Qy 437 SRAYITCASFTSATESTTANTFPLEDHTSENKTDEKKIEERKAQIITEKTSPKTSNPFL 496  
 : | | | : | | : | | |  
 Db 1802 -----DDGSECLDSIGDLSE-----TIQRFN 1824  
 Qy 497 VAVQDSEADYVTTDTLSKVT-----EAAVSNMPEGLTPDLVQEACESELNEATGTKI 548  
 : : | : | : | : : | : : : | |  
 Db 1825 TSIDDPSS---IRRDSFSSISSFGDRQKFRTAIENIRQDLLP----- 1862  
 Qy 549 AYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASV 608  
 : : | : : : | : | : : : | : | : |  
 Db 1863 -FQSSV-----SQYLRSSPNPSQQL-----LVTNLSMDSP-SDLSPNA---- 1898  
 Qy 609 VQPSVSPLEAPPVSYDSI-----KLEPENPPPYEAMNVALKALGTKEGIEKESFNAA 663  
 | | | : : | : | : | : : : :  
 Db 1899 -----PPVGFEATAQFLEKLQEDRPSAEGSID-----SSGFEKVD--HEG 1937  
 Qy 664 VQETEAPYISIACDLIKETKLSTEPSPDF-----SNYSEIAKFEKSVPE 707  
 : | | : | : : : | : : : : | : |  
 Db 1938 LDEFAAPPVH---DPMQKSVFGSLGSDDMKPGSQDDGFVFIERNEANEATLKKNQKMSSH 1994  
 Qy 708 HAEIVEDS-----SPSEPVDLFSDDSIPEVPQTQEEAV----MLMKESLTE 750  
 | : : | : | : | : : : : : : :  
 Db 1995 HNDVIEKNYFNDNAPTAALLESPIAEAEARKLVQDAVESASEYKKQAVDSGDEIGRELLDN 2054

Qy	751	VSETVAQHKE-----ERLSASQPQLGKPYLESFQP	780
Db	2055	VEQKIEQVKEPIVDLSLHKAYDGVGDFVHETVPNAVDDFVREAQKQLPESPVPEKIETPEP	2114
Qy	781	--NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLLSSKEDK-----IKES	830
Db	2115	LVDIHDVTDKVHDEVDNFLRREP-----TPPFETDDVAPLSDDKPQFGNQTPPEDE	2165
Qy	831	ETFSOSSPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGA----DSLPLCL	886
Db	2166	TTFDRKGPLTIPEEVEKAAAAQNND---LDDFDPLVTSNTGAAFGAAVGAAAIVESLTEE	2222
Qy	887	ELPCDLSFKNI-----YPKDEVHVSDEFSENRSSVSKASISPSNVSA-----LEPQTEM	935
Db	2223	EMFGHQKFETVPRPPTPPKD---ISDE-----DVKPSTVNLGPSHHHSHPSSPH	2268
Qy	936	GSIKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGAS	995
Db	2269	HSILKHHG-----DAWIDFKTVPPCVLDVIYWRDAKSAIVLSLA	2308
Qy	996	LFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAIS	1055
Db	2309	LLVLFVLAKYPLLTVVVTYSLLLALGAAAGFRVFKKVEAQIKKTDSEHPFSEILAQDLTLPL	2368
Qy	1056	EELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMMWFTTYVGALFNGLTLLILA	1115
Db	2369	QEKVHAQADVFEHATCIANKLKKLVFVESPLESIKFLVLWSLTYIASWFSGFTLAILG	2428
Qy	1116	LISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLK	1159
Db	2429	LLGVFSVPKVYESNQEAIDPHLATISGHLKNVONIIDEKLPFLR	2472



Qy 28 FVTEPEDEE---DEEEEEDEED-----DEDLEELEVL----- 57  
 | || :| | :| ||| |||: | :|  
 Db 1437 FGTESSEESQKADGNQENQEEEDVVAELNFHPIRQWRDEDVISLQSLKSLVAEVCITDV 1496  
 Qy 58 -----ERKPAAGLSAAAVPPAAAAPL-LDFSSDSVPPAPRGPLPAAPPAAP--ERQPSW 108  
 |: :| | :| :| |||:| | :| || | |  
 Db 1497 DASDVNEQDEESTLKILKVVPPSEPSLLELDFDND--PKVIHVPIPLMEPATMYLEEMVEW 1554  
 Qy 109 -----ERSPAAPAPSLPPAAAVLPS-----KLP-EDDEPPARPPPPPPA 146  
 | : || : : :| ||| ||| | | |  
 Db 1555 IIADAVKEVSEMEVVTSEISEMAPQVSESTCPIPEPLADLKLPVEDDEKTPEPEPVVPG 1614  
 Qy 147 GASPLAEPAAPPSTPAAPKRRGSGSVDETFLALPAASEPVIPS-SAEKIMDLMEQPGNTV 205  
 | | | :| | || :||: | | : | :| |  
 Db 1615 QVQERIPIEVEQAPTIPQRPPRAPKSE----LPKVAKPLDDSKSRVRFAPLNIKLGRY 1670  
 Qy 206 SSGQEDFPSVLLETAASLPSLSPLSTVSKHEGYLGNLSAVSSSEGTIEETLNEASKELP 265  
 | |: | || ||: : : : :|  
 Db 1671 SEEQQK-----ELVESLE--RPLTIIT-----QQKPP 1695  
 Qy 266 ERAT-----NPFVNRDLAEFSELEYSEMSSFKGSPKGESAILVENTKEEVIVRSKDKE 319  
 |: | :| |||: | :| | || : :||: | :|  
 Db 1696 EKPTEDIGALSPLSPNTLAEEYEVPMMDM-QSVPHSP-----QEKQEEIEALSEIE 1746  
 Qy 320 DLVCSAALHSPQES-PVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYADFK---PFEQA 375  
 : : | || | : : :|| | || | :| :| | :|  
 Db 1747 EPQAMKEVEKPVESAPEKDNESELEPE----INE-----PIRRVLVETKIMGPGKSL 1795  
 Qy 376 WEVKDTYEGSRDVLAAARANVESKVDRCLEDSLEQKSLGKDS-----E 418  
 | | : : | : : : :| :| :| :|  
 Db 1796 NEDNDDDDGSECLDSIGDLSERTIQR-FNTSIDDPSIRRDSFSSISSFGDRQKFRTAIE 1854  
 Qy 419 GRNED-----ASFPSTPEPVKDSSRAYITCASFTSATESTT-----ANTFPLLE 462  
 :| :|| | | : :| | : : : || ||  
 Db 1855 NIRQDLLPFQSSVSQYLRSSPNP---SQQLLVTNLSMDSPDLSPNAPPVGFENTAQFLE 1911  
 Qy 463 -----DHTSENKTDEKKIEERKAQII-----TEKTSPKTSNPF 495  
 | : | | : :| | : : :| : :  
 Db 1912 KLQQEDRPSAEGSIDSSGFVKVDHEGLDEFAAPPVHDPMQKSVFGLGSDDMKPGSQDDG 1971  
 Qy 496 LVAVQDSEADYVT-----TDTLSK-----VTEAAVSNMPEGLTPDLVQ 533  
 | : :||: | | :| : : :| :| :|  
 Db 1972 FVFIERNEANEATLKKNQKMSSHNDVIEKNYFNDNAPTAALLESPIAEEARKLVQDAVE 2031  
 Qy 534 EACE-----SELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSVL 588  
 | | : : | : : : | | :||: | | :|  
 Db 2032 SASEYKKQAVDSGDEIGRELLDNVEQKIEQVKEPIVDSLHKAYDGVGDVH-ETVPNAV- 2089  
 Qy 589 PDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKAL 648  
 | || | | ||| :| |  
 Db 2090 DDFVREAE-----KQLPESPVP----- 2106  
 Qy 649 GTKEGIKEPESFNAAVQETEAPYISIACDLIK-----ETKLSTEPSPDFSNYSEIAKF 703  
 | : || | : | : | : | ||:| |  
 Db 2107 ---EKIETPE-----PLVDIHDVTVDKVHDEVNDNFLRREPTPPFE----- 2142



C;Superfamily: Drosophila 576K microtubule-associated protein homolog

Query Match 5.6%; Score 328.5; DB 2; Length 5327;  
Best Local Similarity 22.7%; Pred. No. 4.1e-05;  
Matches 255; Conservative 176; Mismatches 453; Indels 238; Gaps 54;

```
Qy      2 EDIDQSSLVSSSTDSPPRPPPAFKYQFVTEP----EDEEDEEEEEDEEDEDLEEEVL 57
      ::|:| | :| | | :| | :| | | :| | |
Db      3196 DEADKSKEESRRESGAEKSPLASKE--ASRPASVAESIKDEAEKSKEESRRESVAEKSPL 3253

Qy      58 ERKPAAGLSAAA--VPPAAAPLLDFSSDSVPPAPRGPLPAAPPAAP----- 102
      | |: :: | | | :| ||| | :| | : |
Db      3254 PSKEASRPTSAKSVKDEAEKSKEESSRDSV--AEKSPLASKEASRPASVAESVQDEAEK 3311

Qy      103 -----ERQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLA-EPAA 156
      |: | :| | | : | | | | | : : | | | | : | :
Db      3312 SKEESRRESVAEKSPLAYKEASRP-ASVAESIKDEAEKSKEESRRESVAEKSPLASKEAS 3370

Qy      157 PPSTPA-----APKRRGSGSVDETLEFALPAAS-EPVIPSS-AEKIMDLMEQPGNTVSSG 208
      |: | | | : | | | | | | | :| | | : | | :
Db      3371 RPTSVAESVKDEAEKSKEESSRDSVAEKSPLASKEASRPASVAESVQDEAEKSKEESRRE 3430

Qy      209 QEDFPSVLLETAASLPSLSPLSTVSFKEHG-----YLGNL SAVSSSEG TIEETLNEAS 261
      | | | | | : | | | | : | : | | : : | :
Db      3431 SVAEKSPLASKEASRPASVAESVKDDAEKSKEESRRESVAEKSPLASKEASRPASVAESV 3490

Qy      262 KELPERATNPFFVNRDLAEFSELEYSEM--SSFKGSPKGESAILVENTKEEVIVRSKDKE 319
      |: | : : | | | | : | | | | : | : | | : |
Db      3491 KDEAEKSKEESRRESVAEKSPLPSKEASRPTSVAESVKDEA----EKSKEE-----SRRE 3541

Qy      320 DLVCSAALHSPQES-PVGKEDRVV-SPEKTMDFNEMQMSVVAPVREEYADFKPFQAW 377
      : : : | | : | | : | | : : : : | : | |
Db      3542 SVAEKSSLASKEASRPASVAESVKDEAEKSKEESRRESVAEKSPLASKEAS-RPASVAES 3600

Qy      378 VKDTYEGSRDV-----LAARANVESKVDRK--CLEDSLEQKSLGKDSEGRNE----- 422
      ||| | | : : | : | : : | : : : : | | |
Db      3601 VKDEAEKSKEESRRESVAEKSPLPSKEASRPTSVAESVKDEADKSKEESRRESGAEKSP 3660

Qy      423 ---DASFP-STPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTS-----ENKTD- 471
      :| | | | | | : : | | | | : | | | : | |
Db      3661 ASMEASRPTSVAESVKDETEK----SKEESRRESVTEKS-PLPSKEASRPTSVAESVKDE 3715

Qy      472 -EKKIEERKAQIIITEKT---SPKTSNPFVLVAVQDSEADYVTTDTLSKVTEAAVSNMPEG- 526
      || | | : : : | : | : | | : : : | : : |||
Db      3716 AEKSKEESRRESVAEKSPLASKESRP-----ASVAESIKDEAEGTKQESRRESMPESG 3769

Qy      527 -----LTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYP 569
      | | | | : : | : | | : | : | : |
Db      3770 KAESIKGDQSSLASKETSRPDSVVESVKDETEKPEGSAL---DKSQVASRPESVAVSAKD 3826

Qy      570 TAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKL 629
      | | : | | | | | | | : : | : | | |
Db      3827 EKSPLHSRPESVADKS---PDASKEA-----SRSLSVAETASSPIIEGPRSIAD---- 3872

Qy      630 EPENPPPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPS 689
      : : : | | : | | | : : : | : : | : |
Db      3873 -----LSLPLNLTGEAKG-KLP-TLSSPIDVAEGDFL-----EVKAESSPR 3911
```

Qy	690	PDFSNYSEIAKFEKSVPEH--AELVEDSSPESEPVDLF-----SDDSIPEV	733
Db	3912	P--AVLSKPAEFSQPDGTGHTASTPVDEASPVLEEIEVVEQHTTSGVGATGATAETDLLDL	3969
Qy	734	PQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQPNLHSTKDAASNDI	793
Db	3970	TETKSETV--TKQSETTLFETLTSKVESKVEVLESSV-KQVEEKVQTSVKQAETTVDLSL	4026
Qy	794	PTLTKKKIKISLQME-----FNTAIYSNDDLSSKEDKIKESETFSDSSP	838
Db	4027	EQLTKKS--SEQLTEIKSVLDTNISNVTNLFSTAVETIEKKVQDVTEKVIEKAT-----	4078
Qy	839	IEIIDEFPFVSASKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLLEPCDLSFKNIY	898
Db	4079	-EHVSEHVT--TTGESSTETSQEKSSLDLGTFSSELRETHITTVGSPEFTV-----TIC	4128
Qy	899	PKDE--VH-VSDEFSEN-----SSVSKAS-ISPSNVSALEPQTE-----MGSIVKS-	941
Db	4129	ERDEPVLHDIKEEDEEHRFSPSPSDVDKAAIIPPQPMRPLSPREEEVAKIVADVAKVLKSD	4188
Qy	942	KSLT-----KEAEKKLPS--DTEKE-DRSLSAVLSAELS	972
Db	4189	KDITDIIPDFDERQLEELKSTADTEEESDKSTRDEKSLEIS	4230

Db	6634	KKKEAPPAKVPEVPKKPPEEKVPVLIPKKEKPPPAKVPEVPKKPVPEEKVPVPVPPKKVEAP	6693
Qy	99	PA-APE--RQPSWERSPAAPAP---SLPPA-AAVLPSKL-PEDDEP---PAR---PPPPP	144
Db	6694	PAKVPEVPKKPVPEKKVPVPAPKKVEAPPAKVPEVPKKLIPEEKKPTVPVPPKKVEAPPPKV	6753
Qy	145	PAGASPLAEPAAAPPSTPAAPKRRGSGSVDETLFALPAASEPVIIPSSAEKIMDLMEQPGNT	204
Db	6754	PKKREPVPVPVALPQE-----EEVLF-----EEEIVPE--EEVLPEEEEE-----	6790
Qy	205	VSSGQEDFPSVLLETAASLP---SLSPLSTVSFKEHGYLGNLSAVSSSEGTI--EETLNE	259
Db	6791	VLPEEEEE--VLPEEEEEVLPEEEEEIPPEEEVPPEEY-----VPEEEEFVPEEEVLPE	6841
Qy	260	ASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGESA-----ILVENTKEEVIV	313
Db	6842	VKPKVPVPAPVPEIKKKVTEKKVV-----IPKKEEAPPAKVPEVPKKVVEEKRII	6890
Qy	314	RSKDKEDLVCSAALHSPQESPVGKED-----RVVSPEKTMDFNEM-----	354
Db	6891	LPKEEEVLPEVT-EEPEEEPISEEEIPEEPPSIEEVEEVAPPRVPEVIKKAVPEAPTVP	6949
Qy	355	-----QMSVVAPVREEYADFKPFEQAVEVKDITYEGSRDVLAAARANVESK	398
Db	6950	PKKVEAPPAKVSKKIPEEKVPVPVQKKEA---PPAKVPEVPKKVPEKKVLVPKKEAVPPA	7006
Qy	399	VDRKCLEDSE-----LEQKSLGKDSEGRNEDASFPSTPEPVKDSRAYITCA	444
Db	7007	KGRTVLEEKVSVAFRQEVVVKERLELEVVAEEVEEIPEEEEFHEVEE-----YFEEG	7058
Qy	445	SFTSATESTTANTFPLLEDHTSEN-----KTDEKKIEER-----KAQIITEKTSPTK	491
Db	7059	EFHEVEEFIKLEQHRVEEHRVEKVRVIEVFEEAEVEVFEEKPKAPPKGPSEIKIIPPK	7118
Qy	492	SNPFLVAVQDSEADYVTTDTLSKVTEAAV-----SNMPEGLTPDLV-----	532
Db	7119	KPPTKVVPKRKEPPAKVPEVPKKIVVEEKVRVPPEEPRVPPTKVPEVLPPKEVVPEKKVPVP	7178
Qy	533	----QEACESELNEATGTKIAYETKVDL-----VQTSEAIQESLYP	569
Db	7179	PAKKPEAPPPKVPEAP-KEVVPEKKVPVPPPKKPEVPPTKVPEVPKAAVPEKKVPEAIPP	7237
Qy	570	TAQLCPS--FEEAEATPS-----PV-LPDIVMEAPLNSLLPSAGASVVQPSVSPL	617
Db	7238	KPESPPPEVFEEPEESPSAPPKKPEVPPVRVPEVPKEVVPEKKVPAAPPK--KPEVTPVK	7295
Qy	618	AP-----PPVSY-----DSIKLEPENPPP--YEE	639
Db	7296	VPEAPKEVVPEKKVPVPPPKKPEVPPTKVPEVPKVAVPEKKVPEAIPPKPESPPPEVFEE	7355
Qy	640	AMNVALKALGTKEGIKPESEFNAA-----VQETEAPYISIACDLIKETKLSTEP	688
Db	7356	PEEVALLEE-PPAEVVEEPEP--AAPPQVTVPPKNPVPEKKAPAV-----VAKKPELPPVK	7407
Qy	689	SPDFSNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESL	748
Db	7408	VPEVP--KEVVP-EKKVP----LVVPKKPEAPPA-----KVPEVP-----KEVV	7444

Qy	976	VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI	1035
		: ::     ::  :  :    : ::          ::	
Db	27	ILDVIYWRAKKSAILVLSLALLVLFLAKYPLLTVVITYSLLLALGAAAGFRVFKKVEAQI	86
Qy	1036	QKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVL	1095
		: ::        :: : :    :: :   : ::   : :: ::  ::	
Db	87	KKTDSEHPFSEILAQLTLTPQEKVHAQADVFEHATCIANKLKKLVFVESPLESIKFGLV	146
Qy	1096	MWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVIDHYLGLANKSVKDAMAKIQAKI	1155
		:    : :  :         : : ::  :   :     :  : : :    :	
Db	147	LWSLTYIASWFSGFTLAILGLLGVSVPKVYESNQEAIDPHLATISGHLKNVQNIIIDEKL	206
Qy	1156	PGLK 1159	
		:	
Db	207	PFLR 210	

RESULT 9

A47283

calphotin - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000

C;Accession: A47283

R;Ballinger, D.G.; Xue, N.; Harshman, K.D.

Proc. Natl. Acad. Sci. U.S.A. 90, 1536-1540, 1993

A;Title: A *Drosophila* photoreceptor cell-specific protein, calphotin, binds calcium and contains a leucine zipper.

A;Reference number: A47283; MUID:93165730; PMID:8434015

A;Contents: photoreceptor cells

A;Accession: A47283

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-873 <BAL>

A;Cross-references: GB:L05080; NID:g157071; PIDN:AAA28420.1; PID:g157072

A;Note: sequence extracted from NCBI backbone (NCBIN:124958, NCBIIP:124959)

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; von Willebrand factor type C repeat homology

Query Match 5.2%; Score 304.5; DB 2; Length 873;  
Best Local Similarity 21.9%; Pred. No. 3.1e-05;  
Matches 213; Conservative 126; Mismatches 362; Indels 273; Gaps 42;

```

Qy      62 AAGLSAAAVFPAAAAPLLDSSDSVPPAPR--GPLPAAP----PAAP-----ERQPSWER 110
      :| :| | | | | : | :| | | | | | | | | | | | | | | | | | |
Db      11 SAPVAAPVTPSAVAAPVQVVSFAAVAVAPAVVAPAPAPAPTAVTPVAPPPTLASVQPATVT 70

Qy     111 SPAAPAP----SLPPAAAVLPSKLPEDDEPPARPPPPPP-----AGASPLAEPAAPP 158
      | | | | | | : | | :| | : | | | | | | | | | | | | | | | |
Db      71 VP-APAPIAAASVAPVASVAPPVVAAPTTPPAASPVSTPPVAVAQIPVAVSAPVAPPVVAAT 129

Qy     159 STPAAPKRRGSGSVDETLEFALP--AASEP----VIPSSAEKIMDLMEQPGNT---VSSGQ 209
      | | | | | | | : | | | | | | | | | | | | | | | | | | | |
Db     130 PTPVAPI-----PVAAPVIATPPVAASAPTPAAVTPVVSFVIATPPVVPANTTVPVAAAPV 184

Qy     210 EDFPSVLLETAASL-PSLSPLSTVSFKEHGYLEGNSAVSSSEGTEETLNEASKELPERA 268
      | : : | | | | : | | | | | | | | | | | | | | | | | | |
Db     185 AAVPAAVPVAPVLAAPAVAPAVAPVVAETPAPPPVAEIPVA--TIPECVAPLIPEVSVVA 242

Qy     269 TNPVFVNRDLAEFSELEYSEMGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALH 328
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     243 TKPLA-----AAEPVVVAPPATETPVVAPAAA 269

Qy     329 SPQESPVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYADFKPFEQAWVEVKDITYEGSRDV 388
      | | | | | | | : | | | | | | | | | | | | | | | | | | |
Db     270 SPHVS-----VAP-----AVETAVVAPV----- 287

Qy     389 LAARANVESKVDKCLEDSLEQKSL----GKDSEGRNEDASFPSTPEPVKDSSRAYITCA 444
      | : | | | | : | : | | : | | | | | | | | | | | | | |
Db     288 ---SASTEPPVAAATLTAPETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAE 344

Qy     445 SFTSATESTTANTFPLLEDHTSENKTDEKKIEERKAQIITEKTSPKTSNPFLVAVQDSEA 504
      : | : : | | : : | : | | | | | | | | | | | | | | |

```

```

Db      345 TPEVASVAVAETTPPVVPPVAAES-----IPAPVVATTVPVPATLAVTDPD- 389
Qy      505 DYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQ 564
          || :|| :| :| | | | | | | | |
Db      390 -----VTASAVPELPPVVIAPSPVPSA-----VAETPVDLA-----P 420
Qy      565 ESLYP-TAQLCPSF--EEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAP-- 619
          | | |: |: || ||:| : : | |: :| : | |: | |
Db      421 PVLPPVAAEPVPAVVAEETPETPAPASAPVTI-AALD--IPEVAPVIAAPSDAPAEAPSA 477
Qy      620 --PPVSYSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIIACD 677
          | || | | | | | | | | | | | | | | | | | | | | |
Db      478 AAPIVS-----TPPT-----TASVPETTAPPAAVPTE 504
Qy      678 -----LIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAE--LVEDSSPESEPVDFLSD-DS 729
          :: | : | :| :|| : : || || :| | : | || |
Db      505 PIDVSVLSEAAIETPVAPPVEVTTEVAVADVAPPEAAEDLIIPEVEPPAPIPDLLEQTS 564
Qy      730 IPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELGKPYLESFQPNLHSTKDAA 789
          :| | : : : : || :| || : : | :| | :| :| : :
Db      565 VPAVEAAESTSSPIPETSLPPPNEAVA---SPEVAVAPITAPEPIPEP-EPSTATPTPEPI 620
Qy      790 SNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPTFV 849
          : | : : : : : : : : : : : : : : : | | | |
Db      621 PVEAPV-----VIQEAVDAVEVPVTETSTSIPETTVEFPEAV 657
Qy      850 SAKDDSPKLAKEYTDLEV-SDKSEIANIQSGAD----SLPCLLELPCLDSFKNIYPKDEVH 904
          : | | : | : : : : || | | : | :| : | | |
Db      658 AEKVLDPAI----TEAPVTTQEPDVANINDGAPATEITTPAVEI-----VT 699
Qy      905 VSDEFSENRSSVSKASISPSNVSALE-PQTEM--GSIVKSKSLTKEAEKKLPDTEKEDR 961
          : | |: : : | : : | |: || : : :| | | :| : : :
Db      700 AAAEVSDTAIPLIDPPV-PQEIAVAIEPETETKPAEVIVEQS-TIPIEAPVPEVSKYAEP 757
Qy      962 SLSAVLSAELSKTS 975
          :| :||: | :
Db      758 VISEAPAAEVPITA 771

```

RESULT 10

T24583

hypothetical protein T06D8.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T24583

R;Palmer, S.

submitted to the EMBL Data Library, April 1995

A;Reference number: Z19909

A;Accession: T24583

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1829 <WIL>

A;Cross-references: EMBL:Z49130; PIDN:CAA88964.1; GSPDB:GN00020; CESP:T06D8.1

A;Experimental source: clone T06D8

C;Genetics:

A;Gene: CESP:T06D8.1

A;Map position: 2



A;Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3;  
1742/1

Query Match 5.2%; Score 302.5; DB 2; Length 1829;  
Best Local Similarity 21.1%; Pred. No. 0.00011;  
Matches 233; Conservative 157; Mismatches 469; Indels 245; Gaps 38;

```
Qy      2 EDIDQSSLVSSSTDSPPR-----PPPAFKYQFVTEPEDEE-----DEEEE-----E 42
      |:  ::| || : |      |      | || |      : ||      |
Db    212 EETTVAVVESSGEEPASSSTSIPTELSKNDQVTEASGEETITAAATEASEETTTSAVTE 271

Qy      43 DEEEDEDDLEEEVLERKPAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRG---PLPAAPP 99
      || :  ::| :||: | : :|      : | |      |      ||
Db    272 GSGEDTTVAVVESLGEQPAS--SSTSIP-----TELSKDDQVTEASGEETTTAAATE 322

Qy     100 AAPERQPS-----WERSPAAPAPSLPPAAAVLPSKLPEDDE-PPARPPPP 143
      |: | |      | | || |      : :||: ||: |
Db    323 ASEETTTSAVTEGSGEETTVAVVSSGEEPASS---STSIPTELSKDDQVTEASGEET 378

Qy     144 PPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLEFAL-----PAASEPVI PSSAEKIMD 196
      | |: :|      | : || | | : :      ||: | ||: |
Db    379 TTAAATEASE-----ETTTSAVTEGSGE-DTTVAVVESSGEQPASSSTSIPTELSKDDQ 432

Qy     197 LMEQPGN-----TVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEH 237
      : | |      | ||: : :||: | : | :
Db    433 VTEASGEETTTAAATEASEETTTSAVTEGSGEDTTVAVVESSGEQPASSSTSIPT---- 488

Qy     238 GYLGNL SAVSSSEG TIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPK 297
      |      |: : | || | : | | |      | : : : : || :
Db    489 -ELSKDDQVTEASG--EETTTAAATEASEETTTSAVTEGSGEETTV-VAVVSSGEEPAS 544

Qy     298 GESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPV-----GKEDRVVSPEKTMDI 350
      ::| | :||: : : :| :| | :|:      |: | ||: ::|
Db    545 SSTSIPTELSKDDKVTEASGEETTTAAATDASSEETTTSAVTEGSGEETTVAVVSSD- 603

Qy     351 FNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANVESKVDRKCLED SLEQ 410
      | | : | | : | : | |      || |      | | | :
Db    604 -EEPASSSTSIPTELSKDDQVTEASGEETTT-----AAATEASEETTTSAVTEGSGEE 655

Qy     411 KSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTSENKT 470
      :: | | : : || | : | :| | | | :|||      : |: :
Db    656 TTVAVVSSGEEPASSSTSIPTELSKDDKVTEA---SGEETTTAAATDASSEETTTSAV 712

Qy     471 DEKKIEERKAQIITEKTS PK-TSNPFLVAVQDSEADYVTTDTLSKVTEAAAVSNMPEGLTP 529
      | ||      : | : : | : : : | | | : : | || : | |
Db    713 TEGSGEETTVAVVSSDEEPASSSTSIPTELSKDDQVTEASGEETTTAAATEASEETTT 772

Qy     530 DLVQEACESELN-----EATG-----TKIAYETKVDLVQTSEAIQESLYPTAQLCPSF 577
      | | |      | : : | | | | | :|| |      | :
Db    773 SAVTEGSGEETTVAVVSSGEEPASSSTSIPTELSKD-DQVTEASGEE---TTAAATE 828

Qy     578 EEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPY 637
      | | | |      : : : : | : | ||: |      | :|
Db    829 ASEETTTSAVTEGSGEDTTVAVVSSGEE---QPASSSTSIPTELS----- 871

Qy     638 EEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFS NYSE 697
      | : : | : :||      : | : :|| |      :

```

Db 872 -----KDDQVTEASGEETTT---AAATEASEETTTSAVTEGSGEETTV 911

Qy 698 IAKFEKSVPEHAELVEDSSPESEPVDFSDDSIPEV---PQTQEEAVMLMKESLTVSET 754  
 :| | | | | | | | | | :| | | :| | | :|

Db 912 VAVVSSGEEPA-----SSSTSIPTELSKDDQVTEASGEETTTAAATEASSEETTTSAVT 966

Qy 755 VAQHKEERLSASPQELGKPYLESFQP-----NLHSTKDAASN----- 791  
 :| | | : | : | | : | | : | | : |

Db 967 EGSGEETTTSAVTEGSGEETTTSAVPEGENSTTEAPAFVTGSEIEIPSSEESSSTTHDP 1026

Qy 792 DIPTLTKKKIKISLQMEEFNTAIYSNDDLSSKEDIKESSET-FSDSSPIEIIIDEFPTFVS 850  
 || :| | :| :| : | : : : | | :| | : | | | | :

Db 1027 SIPVITPKPSVSSTIENVMSKTSSEE---AAEKKIIGEHQTGKDDAGKEDEDNMPAFVT 1083

Qy 851 A-----KDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCD 891  
 | :| : | :| | :| :| :

Db 1084 ANPAGTSTTESAENVSTSTGEEDENIKMAKELGKQFAADLAKLA----- 1126

Qy 892 LSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIKSKSLTKEAEKK 951  
 || | : : | : : | | : | | : : | | : : | | :

Db 1127 -----AKDGVNLT-ETADAKDSGETAHVEDEQVSSSTE--SSIGSEETTTTVNKETTEE 1176

Qy 952 LPSDTEKEDRSLSAVLSAELSKTS 975  
 : : | | : : | | :

Db 1177 HHEASGEEDDAPAFVTGAPTDSTT 1200

# RESULT 11

A47282

calcium-binding protein calphotin - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C;Accession: A47282

R;Martin, J.H.; Benzer, S.; Rudnicka, M.; Miller, C.A.

Proc. Natl. Acad. Sci. U.S.A. 90, 1531-1535, 1993

A;Title: Calphotin: a *Drosophila* photoreceptor cell calcium-binding protein.

A;Reference number: A47282; MUID:93165729; PMID:8094559

A;Accession: A47282

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-865 <MAR>

A;Cross-references: GB:L02111; NID:g157031; PIDN:AAA28405.1; PID:g157032

A;Experimental source: photoreceptor cells

A;Note: sequence extracted from NCBI backbone (NCBIN:124955, NCBIP:124956)

C;Genetics:

A;Gene: FlyBase:Cpn

A;Cross-references: FlyBase:FBgn0010218

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; von Willebrand factor type C repeat homology

C;Keywords: calcium binding

Query Match 5.1%; Score 295.5; DB 2; Length 865;

Best Local Similarity 21.3%; Pred. No. 7.2e-05;

Matches 217; Conservative 127; Mismatches 379; Indels 295; Gaps 41;

Qy 62 AAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPER-----QPSWERSPAAPAP 117  
 :| :| | | | | :| | | | :| | | | | | | | | |

Db 11 SAPVAAPVTPSAVAAPVQVVS PAAVAPAPAAPIAVTPVAPPPTLASVQPATVTIP-APAP 69  
 Qy 118 ----SLPPAAAVLPSKLPEDDEPPARPPPPPPAGA-----SPLAEPAAPSTP----- 161  
 | : | : | | : | | | | | : | : | | | |  
 Db 70 IAAASVTPVASVAPPVVAAPTPPAASPVSTPVAVAQIPVAVSAPVAPPVVAATPTPVVQIP 129  
 Qy 162 -AAPKRRGSGSVDETFLFALP--AASEP----VIPSSAEKIMDLMEQPGNT--VSSGQED 211  
 | | | : | | | | | | : | | | | : :  
 Db 130 VAAP-----VIATPPVAASAPTPAAVTPVISPVIASPPVVPANTTVPVAAPVAA 178  
 Qy 212 FPSVLLETAASL-PSLSPLSTVSVFKEHGYLGNL SAVSSSEG TIEETLNEASKELPERATN 270  
 | : : | | | : : | | : : : | | | : | : | |  
 Db 179 VPAAPVVPVAPVLAPAVAPAVAPVVAETPAPPPVAEIPVA--TIPECVAPLIPEVSVVATK 236  
 Qy 271 PFVNRDLAEFSELEYSEMGS SFKSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSP 330  
 | | : | | | : | | : | | |  
 Db 237 PLA-----AAEPVVVAPPATETPVVAPAAASP 263  
 Qy 331 QESPVGKEDRVVSPEKTMDFNEMQMSV VAPVREEYADFKPF EQAWEVKDTYEGSRDVL A 390  
 | | : | : : | | | |  
 Db 264 HVS-----VAP-----AVETAVVAPV----- 279  
 Qy 391 ARANVESKVDRKCLED SLEQKSL----GKDSEGRNEDASFPSTPEPVKDSSRAYITCAS F 446  
 | : | | | : | : | : | : | | | : : :  
 Db 280 -SASTEPPVAAATLT TAPETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAETP 338  
 Qy 447 TSATESTTANTFP LLEDHTSENKTDEKKIEERKAQIITEKTS PKTSNPFLVAVQDSEADY 506  
 | : : | | : : : : | | | : | | :  
 Db 339 EVASVAVAETTPPVVPPVAAES-----IPAPVVATTPVPATLAVTDPD--- 381  
 Qy 507 VTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATG TKIAYETKVDLVQTSEAIQES 566  
 | | : | | : | | | | | | | |  
 Db 382 -----VTASAVPELPPVIA P SPVPSA-----VAETPVDLA-----PPV 414  
 Qy 567 LYP-TAQLCPSF--EEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAP---- 619  
 | | | : | : | | | : | : | : | : | : | | | |  
 Db 415 LPPVAAEPVPAVVAEETPETPAPASAPVTI-AALD--IPEVAPVIAAPSDAPAEAPSAAA 471  
 Qy 620 PPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIACD-- 677  
 | | | | | | | : | | | | : : :  
 Db 472 PIVS-----TPPT-----TASVPETTAPPAAVPTEPI 498  
 Qy 678 ---LIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAE--LVEDSSPESEPVDFLSD-DSIP 731  
 : : | : | : | : | : : | | | : : | | : | |  
 Db 499 DVSVLSEAAIETPVAPPVEVTTEVAVADVAPPEAAADLIEPVEPPAPIPDLLEQTTSVP 558  
 Qy 732 EVPQTQEEAVMLMKESL TEVSETVAQHKEERLSAS PQELGKPYLESFQPNLHSTKDAASN 791  
 | : : : : | | : | | : : | : | | : | : :  
 Db 559 AVEAAESTSSPIPETSLPPPNEAVA---SPEVAVAPITAPEPIPEP-EP SLATPTEPIPV 614  
 Qy 792 DIPTLT KKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIEIDEFPTFVSA 851  
 : | : : | : : | : : | : : | | | | :  
 Db 615 EAPV-----VIQEA VDAVEVPVTETSTSI PETTVEFPEAVAE 651  
 Qy 852 KDDSPKLAKEYTDLEV-SDKSEIANIQSGAD----SLPCLEL-----PCDLSFKNIYP- 899  
 | | : | : | : : : | | : | : : | : : | |  
 Db 652 KVLDPAI----TEAPVTTQEPDVANINDGAPATEITTPAVEIVTAAAEVSDIAIPVIDPP 707

Qy 900 -KDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMG----- 936  
 | : | : | : : : | : | :  
 Db 708 VPQEIAVA-EIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAAEVPITAGDN 766  
 Qy 937 ----SIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGV 990  
 | : | : : ||| : | : | : : || ||: | |  
 Db 767 PDNTSVGISEVVPTIAEKPVEEVPTSEIPEQSSSPSDSVPVAKITPLL--RDLQTTDV 822

# RESULT 12

T34513

hypothetical protein ZK783.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T34513

R;Favello, A.; Vaudin, M.

submitted to the EMBL Data Library, August 1994

A;Description: The sequence of *C. elegans* cosmid ZK783.

A;Reference number: Z21536

A;Accession: T34513

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-3507 <FAV>

A;Cross-references: EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1

A;Experimental source: strain Bristol N2; clone ZK783

C;Genetics:

A;Gene: CESP:ZK783.1

A;Map position: 3

A;Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1;  
 1409/2; 1450/1; 1760/1; 1898/1; 2153/1; 2628/1; 2681/1; 2899/1; 2946/1; 3088/3;  
 3184/1; 3283/1; 3346/3; 3365/3; 3484/3; 3504/1

Query Match 5.0%; Score 292; DB 2; Length 3507;  
 Best Local Similarity 20.6%; Pred. No. 0.00072;  
 Matches 226; Conservative 146; Mismatches 491; Indels 236; Gaps 37;

Qy 11 SSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDL--EELEVLERKPAAGLSAA 68  
 |||::| || : :|| : :|| | | : || | : |  
 Db 2043 SSSSEAPLTSSPATTEVITESSVKSTTPKEESSSEITVKLSSKSPEVTESSVKSSPSTP 2102  
 Qy 69 AVPPAAAAPLLDFSSDSVPPAPRGPLPAAPP----AAPERQPSWERSPA-----APAPSL 119  
 : : : :| | : | : : | : | : || | : ||  
 Db 2103 STTSQSVTSTVPETSKSTVLSSEAPVTSTSPTEVHTSSETKPSLSASSTTGDNTSTTPST 2162  
 Qy 120 PPAAAVLPSKLPEDDEPPARPPPPPPAGASP-LAEP-----AAPPSTPAAPKRRGSGSV 172  
 | : | : || | | : || ::| | || | : ||  
 Db 2163 SSLASVKSTSAPEGTS--ASVAPVKLSSLSPDVSPSTKTTFDATESSTVQASETSSGTSV 2220  
 Qy 173 DET-----LFALPAASEPVI-PSSAEKIMDLMEQPGNTVSSGQEDFP----SV 215  
 | : : :|| | | : : ||| :| ||| | |  
 Db 2221 KSTSEPEHVTKLSITSSNPSSSVPVTSKSTPTVPESTEQPTSTTPSGQSLTPMNSNSE 2280  
 Qy 216 LLETA---ASLPSLSPLSTVFSKEHGYLGNLSAVSSSEGTIEETLNEASKELPERAT--N 270  
 :| | : |||| : | | | : : : | :|| | | :  
 Db 2281 VLTTSEPHVLSSSLSPDVQSSTTPNNLSESSTVETPKTSSEVSLNSEEPSTTEAPTTL 2340

Qy	271	PFV----NRDLAEFSEL---EYSEMGSFSGSPKG-----ESAILVENTKEEVIVRSKD	317
Db	2341	PDILSTTTNNLSQSSTVSTEDRSEISSENSEKPTSAPELVTSVTHVASSSPDVPTESSE	2400
Qy	318	KEDLVCSAALHSPQ-----ESPVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYADFK	370
Db	2401	PDDLTGSSTENIPEASSKQTISSTPTPDTTTASEEPTKSTSMSPDLSTTSNVLSSESSTTP	2460
Qy	371	PFEQAWEVKDTYEGSRDVLAAANVESKVDKRCLEDSLEQKSLGKDSEGRNEDASFPSTP	430
Db	2461	ESSSKSPVSSSTEG---ISVVTSTEFKSKVPESTISSVLE-----EDLT-KTTP	2504
Qy	431	EPVKDSSRAYITCASFTSAT-----ESTTANTFPLLEDHTSENKTDEKKIEE	477
Db	2505	SPILEET----TTASETSEPLTEDSLTVSVRIHELTTSSENVPKESSESTTTSSESSKPSQ	2560
Qy	478	RKAQIITEKTS PKTSNPFVLAVQDSEADYVTTDTLSKVTEAAVSNMPEGL-----TPDLV	532
Db	2561	EPAGILTSTVVVPTSSVSLITASEIEA--ITSNTPFKQGRTPITTS PKSLVKSTTSPSTV	2618
Qy	533	--QEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLDP	590
Db	2619	TSSEPSESTKRITTVSTTVSTTTPTTEETTTSESLILTAAPSK---PTESTESSEAPTTP-	2674
Qy	591	IVMEAPLNSLLPSAGASVVQPSVSPLEAPPVSYDSIKLEPENPPPYEAMNVALKALGT	650
Db	2675	-----AKTSETKPS-----NVSSTSRKS	2692
Qy	651	KEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAE	710
Db	2693	TENVETSTSQSGSLESS-----TMSSTTSSEPETNAPAVTVSSEASSTTLEE	2738
Qy	711	LVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKE-ERLSASPQE	769
Db	2739	NSSTSSPTSSSEASVKLSSLFPE--SITSEAVTVSSRAPAEITMSSESHREISTVSSEPSE	2796
Qy	770	LGKPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLLSS-----	822
Db	2797	PEIPLSTTVSPNVVTASSIPSEE-PILSSVTSSSTPRVRLITG--TPDDLIVSVTVPSHG	2853
Qy	823	-KEDKIKESETFSDS-SPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGA	880
Db	2854	NRRQNITASSVPSNSTSPIILPSESLLTPQPPPTTTTTAKPAT-----TSGK	2900
Qy	881	DSLPCLELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVK	940
Db	2901	RGPPSIQPPAEM-----FTTPAPPPPSNGGYGEE-----	2929
Qy	941	SKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLL	1000
Db	2930	----TNQEEEQVTSTTTTEAPSLCSTVTCHSLATCE-----QSTGVCICRDGFIGD	2976
Qy	1001	SLTVFSIVSVTAYIALALL	1019
Db	2977	GTTACSKKSTADCISLPSL	2995

RESULT 13

A56577

microtubule-associated protein MAP 1B - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Feb-1997

C;Accession: A56577

R;Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.

Eur. J. Cell Biol. 57, 66-74, 1992

A;Title: Identification of two distinct microtubule binding domains on recombinant rat MAP 1B.

A;Reference number: A56577; MUID:92347374; PMID:1639092

A;Accession: A56577

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2364 <ZAU>

A;Cross-references: GB:X60550

A;Experimental source: brain

A;Note: nucleotide sequence not given; conceptual translation not complete

C;Superfamily: microtubule-associated protein MAP1B

Query Match 5.0%; Score 291.5; DB 2; Length 2364;  
Best Local Similarity 20.0%; Pred. No. 0.00043;  
Matches 220; Conservative 167; Mismatches 456; Indels 255; Gaps 44;

```

Qy      30 TEPEDDEEEEEDEEEDDEDLEELVLERKPAAGLSAAAVPPAAAPLLDFSSDSVPPA 89
      :| | ||:|: || | : : : | | | | : : | :
Db      913 SEEEGEEEEEDKAEDAREEDHEPDKTE-----AEDYVMAVVDKAAEAGVTEDQYDFL--- 963

Qy      90 PRGPLPAAPPAAPERQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGAS 149
      | :| | :|| ||:: | | | | |
Db      964 -----GTPAKQ-----PGVQSPSREPASSIHDETLPGGSESEAT-----AS 999

Qy     150 PLAEPAAPPSTPAAPKRRGSGSVDETLFALPAASEPVIPSSAEKIMDLMEQPGN---TVS 206
      | | | | : :|| | :| | |
Db     1000 DEENREDQPEEFTAT----SGYTQST---IEISSEPTPMDEMSTPRDVMTDETNNETES 1052

Qy     207 SGQE-----DFPSVLLETAASLP---SLSPLS---TVSFKEHGYLGNL SAVSSSEG TIE 254
      || : | | | | ; || | : | : | |
Db     1053 PSQEFVNITKYESSLYSQEYSKPVVASFNGLS DGSKTDATDGRDYNASASTISPPSSMEE 1112

Qy     255 ETLINEAS-----KELPERATNPFVNRDLAEF--SELEYSEMG--- 289
      : ::: |:: : | : | : | : : |
Db     1113 DKFSKSALRDAYRPEETDVKTGAELDIKDVSDERLSPAKSPSLSPSPSPSPSPSPSPSPSPSP 1172

Qy     290 -----SSFKGSPKGESAILVENTKEEVIVR-----SKDKEDL-----VCSAALHSP 330
      : | | :||: :| : :| | ::: | | :| | :|
Db     1173 VNFSLTPNEIKASAEGEATAVSPGVTQAVVEEHCASPEEKTLEVVSQSVTGSAGHTP 1232

Qy     331 -QESFVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYADFK-----PFEQAWVEVKD 380
      :|| :: : | | : : || | : : | | : : | |
Db     1233 YYQSPTDEKSSHLPTTEVT-----ENAQAVPVSFTEAKDENERSSISPMDE--PVPD 1283

Qy     381 TYEGSRDVLAAARANVESKVDRKCLED--SLEQKSLGKDS----EGRNEDASFPSTPEPVK 434
      : ||: : || | : |:: | ||: | | ||
Db     1284 SESPIEKVLSPLRSPPLIGSESAYEDFLSADDKALGRRSESPFEGKNGKQGFSDKESPV 1343

Qy     435 DSSRAYITCASFTSATESTTANTFPLLEDHTSENKT-----DEKKI----- 475

```

```

      |      :|      :      |      |      | : || : | |      ||:|:
Db      1344 D-----LTSDLYQDKQEEKRAGFIPIKEDFSPEKKASDAEIMSSQSALALDERKLGGDGS 1398
Qy      476 -----EERKAQIITEKTS PKTSPKTSNPFLLVAVQDSEADYVTTDTLSKVTEAAV-- 520
      | : | | | : :|: | :|: | : :| :|:
Db      1399 PTQVDVSQFGSFKEDTKMSISEGTVSDKSATPVDEGAEDT---YSHMEGVASVSTASVAT 1455
Qy      521 SNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQES-LYPTAQLCP---S 576
      | : || | | | : :|: | :|: | :|| | : :| : :| | |
Db      1456 SSFPEPTTDD-VSPSLHAEVGS PHSTEVDDSLSVSVVQTPTTTFQETEMSPSKEECPRPMS 1514
Qy      577 FEEAEATP-----SPVLPDIVMEAPLNSLL-----PSAGASVVQ 610
      : :| | : : | :| :|| | :| :|| :
Db      1515 ISPPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSLAMDFSRQSPDHPTVGAGMLH 1574
Qy      611 PSVSPLEAPPPVSYD-----SIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFN 661
      : : | | | | :| :| | | : : : : : : | : :
Db      1575 ITEN---GPTEVDYSPSDIQDSSLSHKIPPTTEEPSYTQDNDLS-ELISVSQVEASPSTSS 1630
Qy      662 AAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESEP 721
      | | | :| | | | :| :| : | :| | :| :| :|
Db      1631 AHTPS-----QIASPLQEDTLSDVVPFRDMSLYASLASEKVQSLEGEKL---SPKSDI 1680
Qy      722 VDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELGKPYL-ESFQP 780
      | :| | : : ||| | :| : : :| | | :| :|
Db      1681 SPLTPRESSPTYSPGFSDSTSGAKES-TAAYQTSSSPIDAAAAPYGFSSMLFDTMQH 1739
Qy      781 NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE 840
      :| :| : : : :| : || | : | : : :| | | :|
Db      1740 HLALSRDLTTSSV----EKDNGGKTPGDFNYAYQKPESTTESPDEEDYDYE--SHEKTIQ 1793
Qy      841 IIDEFPTFVSAKDDSPK----LAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKN 896
      | : : : | | :| :| : :| | | : :|
Db      1794 AHDVGGYGYEKTERTIKSPCDSGYSYETIEKTTKTP----EDGGYS-----CEITEKT 1842
Qy      897 IYPKDEVHVSDEFSENRSSVSKAS-----ISPSNVSALEPQTEMGSIVKSKSL 944
      :| | | | : :| : : : : :| :| :|
Db      1843 TRTPEEGGYSEISEKTRTPEVSGYTYEKTERRRLDDISNGYDDTEDGGHTLGDCSY 1902
Qy      945 TKEAEKKLPSDTEKEDRS 962
      : | :| :| | | |
Db      1903 SYETTEKITSFPESESYS 1920

```

# RESULT 14

T19431

hypothetical protein C25A1.10 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000

C;Accession: T19431

R;Mortimore, B.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19124

A;Accession: T19431

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-971 <WIL>

A;Cross-references: EMBL:Z81038; PIDN:CAB02755.1; GSPDB:GN00019; CESP:C25A1.10  
A;Experimental source: clone C25A1  
C;Genetics:  
A;Gene: CESP:C25A1.10  
A;Map position: 1  
A;Introns: 38/3; 92/3; 201/3; 919/3  
C;Superfamily: nucleolus-cytoplasm shuttle phosphoprotein

Query Match 4.9%; Score 288.5; DB 2; Length 971;  
Best Local Similarity 20.6%; Pred. No. 0.00016;  
Matches 198; Conservative 127; Mismatches 391; Indels 247; Gaps 35;

```

Qy      7  SSLVSSSTDSPRRPPPAFKYQFVTEP---EDEEDEEEEEDEEDEDLEEEVLERKPAA 63
      || | | | ||: || : : : : : | : || :: :: | |
Db     147  SSSDSDSDDEPPKKAPAVTTKVAPKPMAKKQDTSDDSDSDSDSDGSKKANPVKVTTPVA 206

Qy     64  GLSAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAP-----ERQPSWE 109
      : | | ||: | | | ||: | | | | | | | | | | | | | | | | | | | | | | | |
Db     207  NVLQKVVAKKAASSSSDSSDDEKKPAAK-PTPAKPTPKPVVKAESSSDSSDDEKKPVAK 265

Qy     110  RSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGS 169
      : || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     266  PPAKATPK-PAKKADSSSDSSDDEAPAK-KTPAKAAPKPVAKKAESSSDSSDDEKK-- 321

Qy     170  GSVDETLFALPAASEPVIPISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPL 229
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     322  -----PAAK----PTPAKATPKPVAKKAESSSDSSDDEKKPVAKPAKATPKPV 367

Qy     230  STVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSE-- 287
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     368  AK-----KAESSSDSSDDEK-KPAAKPTPAKATPKPVAKKAESSSDSSDDEKK 414

Qy     288  ---MGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSP 344
      : | | : || | | | | | | | | | | | | | | | | | | | | | | | | |
Db     415  PPAKPTSAKATPK-----PAAKKAD---SSSDSSDDEAPAKKTPAKAAP 455

Qy     345  E---KTMDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRD-----VLAARANV 395
      : | : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db     456  KPASKKAESSSDSSDDEKPAKSTPAKITPKPTAKKVASSSSDSSDDEKKPAKPTPANA 515

Qy     396  ESKVDRKCLEDL-----EQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSAT 450
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     516  TPKPVAKKAESSSDSSDDEKKPVAKPTSAK-----ATPKPAAKKADLSSDFSDDEAPA 568

Qy     451  ESTTANTFPLLEDHTSENKTDEKKIEERKAQIITEKTSPK-TSNPFLVAVQDSEADYVTT 509
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     569  KKTPAKAAPKPAKKAESSSDSSDDEKPAKSTPAKTPKPTAKKAASSSSDSSDD--EK 626

Qy     510  DTLISKVTEAAVSNMPEGLTPDLVQACESELN-----EATGTKIAYETKVDLVQTSEAI 563
      :: | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db     627  KPVAKPTSAKATPKPAKKADSSSDSSDDEAPAKKTPVKPTPVKIVAK-KVD---SSSDSS 683

Qy     564  QESLYPTAQLCPSFEEAEATPSVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVS 623
      : || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     684  DDEKKPT-----KATPVKVTP-----KSVTKKAAAS-----S 710

Qy     624  YDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIIACDLIKETK 683

```



Db	711	SDS--SDDEKKPVVKQTPNVV-----PKKEKAASSSDSS-----DDEKK	748
Qy	684	LSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESEPVDFSDDSIPEVPQTQEEAVML	743
Db	749	PTAKPTP-----KATPKQSAKKADSSDDS-----SDDEAPA-----	779
Qy	744	MKESLTEVSETVAQHKEERLSASPQELGKPYLESFQPNLHSTKDAASNDIPTLTKEKIS	803
Db	780	-----KKTPAKSTPAKTAVKKEASS	799
Qy	804	LQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDFPTFVSAKDDS--PKLAKE	861
Db	800	-----SSDSSDDEKTKKKSATTPAKSTPKTALKKAESSDSSDDDEDLPKPSKA	848
Qy	862	YTDL-EVSDKSEIANIQSGADSLPCLEL-PCDLSFKNIYPKDEVHVSDEFSENRSSVSKA	919
Db	849	VTPRPQRADSEESAETEESSRTPALKAKPLATSTEKAVYENRKRKSSPF--RRVQMTKD	906
Qy	920	SIS 922	
Db	907	SVS 909	

# RESULT 15

## QRMSP1

microtubule-associated protein MAP1B - mouse

N;Alternate names: microtubule-associated protein MAP1(X); microtubule-associated protein MAP1.2; microtubule-associated protein MAP5

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 01-Sep-2000

C;Accession: S07549; S44387; A33645

R;Noble, M.; Lewis, S.A.; Cowan, N.J.

J. Cell Biol. 109, 3367-3376, 1989

A;Title: The microtubule binding domain of microtubule-associated protein MAP1B contains a repeated sequence motif unrelated to that of MAP2 and tau.

A;Reference number: A33645; MUID:90094539; PMID:2480963

A;Accession: S07549

A;Molecule type: mRNA

A;Residues: 1-2464 <NOB>

A;Cross-references: EMBL:X51396; NID:g52999; PIDN:CAA35761.1; PID:g53000

R;Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.

Arch. Biochem. Biophys. 310, 428-432, 1994

A;Title: Binding of heat-shock protein 70 (hsp70) to tubulin.

A;Reference number: S44387; MUID:94234720; PMID:8179328

A;Accession: S44387

A;Status: preliminary

A;Molecule type: protein

A;Residues: 653-663, 'IC' <SAN>

C;Superfamily: microtubule-associated protein MAP1B

C;Keywords: microtubule binding; phosphoprotein; tandem repeat

F;589-786/Domain: microtubule binding #status experimental <MTB>

F;589-592, 639-642, 649-652, 655-658, 660-663, 668-671, 674-677, 679-682, 683-686, 687-690, 691-694, 695-698, 699-702, 708-711, 712-715, 716-719, 720-723, 727-730, 758-761, 764-767, 783-786/Region: 4-residue repeats (K/R-K-E/D-X)

F;1861-2064/Region: 17-residue repeats

F;91,116,351,888,1124,1153,1168,1208,1662,1877,1918,2003,2030,2054,2083/Binding site: phosphate (Ser) (covalent) #status predicted  
F;147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr) (covalent) #status predicted  
F;1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 4.9%; Score 284; DB 1; Length 2464;  
Best Local Similarity 20.7%; Pred. No. 0.00093;  
Matches 233; Conservative 159; Mismatches 424; Indels 312; Gaps 51;

```
Qy      31 EPEDEEDEEEEEDEEED-----DEDLEELE-----VLERKPAAG-----LSAAAVP 71
      | | | : | | | | : | | : | | | : | | : |
Db    1009 EAEQSEEEGEEEDKAEDAREEGYEPDKTEAEDYVMAVADKAAEAGVTEEQYGYLGTSAKQ 1068

Qy      72 PAAAPLLDFSS----DSVPPAPRGPLPAAPPAAPERQP-----SWERSPAAPA 116
      | : | : : | : : | | : | | | : |
Db    1069 PGIQSPSREPASSIHDETLPGGSESEATASDEENREDQPEEFTATSGYTQSTIEISSEPT 1128

Qy     117 P---SLPPAAAVLPSKLPEDDEPPARP-----PPPPPAGASPLAEP-- 155
      | : | : | : | | : : | | : | : |
Db    1129 PMDEMSTPRDVMSDETNNEETESPSQEFVNITKYESSLYSQEYSKPAVASFNGLSEGSKT 1188

Qy     156 -----APPSTPAAPKRRGSGSVDETLFALPA-----ASEPVI PSSAEKIMDLMEQ 200
      | | | : | | : | | : | : | : | : | : |
Db    1189 DATDGKDYNASASTISPP-----SSMEEDKFSKLSALRDAYCSEEKELKASAE--LDIKDV 1241

Qy     201 PGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYLGNL SAVSSSEG TIEETLNEA 260
      : | | | | | | | | | | | | | | | |
Db    1242 SDERLS-----PAKSPSLSP-----SPPSPIEKT---- 1265

Qy     261 SKELPERATNPFVNRDLAEFS----ELEYSEMSSFGKSPKGESAILVEN--TKEEVIVR 314
      | | | : | | | | : | | : | | | : | : | : |
Db    1266 --PLGERSVN-----FSLTPNEIKVSAEGEARSVSPGVTQAVVEHCASPEEKTLE 1314

Qy     315 SKDKEDLVCSAALHSP-QESPVGKEDRVVSPEKTMDFNEMQMSVVA-PVREEYADFK-- 370
      | : | | : | : | : | : | : | : | : | : |
Db    1315 VVSPSQSVTGSAGHTPPYQSP-----TDEKSSHLPTEVSENAQAVPVSEFSEAKDE 1366

Qy     371 -----PFEQAWEVKDTYEGSRDVLAAANVESKVDRKCLED--SLEQKSLGKDS---- 417
      | : : | | : | | : : | | | : | | : |
Db    1367 NERASLSPMDE--PVPDSESPVEKVLSPRLSPPLLGSESPYEDFLSADSKVLGRRSESPF 1424

Qy     418 EGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKT----- 470
      | | : | | | | : : : | : | : | | |
Db    1425 EGKNGKQGFPDRESPVSDLT----STGLYQDKQEEKSTGFIPIKEDFGPEKKTSDVETMS 1480

Qy     471 -----DEKKI-----EERKAQIITEKTS PKTSNPFLVAVQDSEAD 505
      | | : : | : | | | : | | | : |
Db    1481 SQSALALDERKLGGDVSPQTQIDVSQFGSKEDTKMSISEGTVSDKSATPVDEGV--AEDT 1538

Qy     506 YVTTDTLSKVTEAAV--SNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAI 563
      | : : : | : | | | : : | : | : | : |
Db    1539 YSHMEGVASVSTASVATSSFPEPTTDD-VSPSLHAEVGS PHSTEVD DSLSVSVVQTPTTF 1597

Qy     564 QES-LYPTAQLCP---SFEEAEATP-----SPVLDPDIMEAPLNSLL-- 601
      | | : : | : | | : : | : : | : | : |
Db    1598 QETEMSPSKEECPRPMSISPPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSFAMD 1657
```

Qy 602 -----PSAGASVVQPSVSPLEAPPPVSYD-----SIKLEPENPPPYEEAMNVA 644  
 | : | | | : : : | | | | : | | : : :  
 Db 1658 FSRQSPDHPTLGASVLHITEN---GPTEVDYSPCDIQDSSLSHKIPPTTEEPSYTQDNDLS 1714

Qy 645 LKALGTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKS 704  
 : : : | : : | | | : : | : | : : |  
 Db 1715 -ELISVSQVEASPSTSSAHTPS-----QIASPLQEDTLSDVVPREMSLYASLA----- 1762

Qy 705 VPEHAELVEDSSPESEPVDFSDDSIPEVPQTQEEAVML----MKESLTVSETVAQHKE 760  
 : | : | | | : | | | | : | : : | | | |  
 Db 1763 ----SEKVQ--SLEGEKLSPKSDIS----PLTPRESSPLYSPGFSDSTSAAKETAAAH-- 1810

Qy 761 ERLSASP---QELGKPY-----LESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFN 810  
 : | : | : | : : : : : : : : : :  
 Db 1811 -QASSSPPIDAATAEPYGFRSSMLFDTMQHHLALNRDLTTSSV----EKDSGGKTPGDFN 1865

Qy 811 TAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPTFVSAKDDSPK----LAKEYTDLE 866  
 | : : | : : : | | | : : : | : |  
 Db 1866 YAYQKPENAAGSPDEEDYDYE--SQEKTIRTHDVVRYYYEKTERTIKSPCDSGYSYETIE 1923

Qy 867 VSDKSEIANIQSGADSLPCLELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKAS----- 920  
 : | : : | | : : | : | | | : : |  
 Db 1924 KTTKTP----EDGG-----YTCEITEKTTTRTPEEGGYSEISEKTTTRTPEVSGYTYEK 1972

Qy 921 -----ISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKEDRS 962  
 : : : : | : | : | : | | | |  
 Db 1973 TERSRRLDDISNGYDDTEDGGHTLGDCSYSYETTEKITSFPESESYS 2020

Search completed: September 29, 2004, 18:15:50  
 Job time : 36.465 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:16:15 ; Search time 124.84 Seconds  
(without alignments)  
2997.869 Million cell updates/sec

Title: US-09-830-972-2  
Perfect score: 5848  
Sequence: 1 MEDIDQSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRRKAD 1163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		Query			Description	
No.	Score	Match	Length	ID		

1	5848	100.0	1163	9	US-09-893-348-18	Sequence 18, Appl
2	5307	90.7	1163	12	US-10-267-502-431	Sequence 431, App
3	4403.5	75.3	1192	9	US-09-758-140-6	Sequence 6, Appli
4	4403.5	75.3	1192	9	US-09-972-599A-6	Sequence 6, Appli
5	4403.5	75.3	1192	14	US-10-060-036-71	Sequence 71, Appl
6	4398.5	75.2	1192	9	US-09-789-386-2	Sequence 2, Appli
7	4398.5	75.2	1192	9	US-09-893-348-23	Sequence 23, Appl
8	4398.5	75.2	1192	12	US-10-267-502-429	Sequence 429, App
9	4398.5	75.2	1192	16	US-10-327-213-9	Sequence 9, Appli
10	4398.5	75.2	1192	16	US-10-466-258-9	Sequence 9, Appli
11	4389.5	75.1	1192	12	US-10-408-967-7	Sequence 7, Appli
12	1513	25.9	379	14	US-10-205-194-164	Sequence 164, App
13	1411.5	24.1	360	9	US-09-893-348-20	Sequence 20, Appl
14	1191	20.4	373	9	US-09-789-386-6	Sequence 6, Appli
15	1191	20.4	373	9	US-09-765-205-6	Sequence 6, Appli
16	1191	20.4	373	9	US-09-893-348-24	Sequence 24, Appl
17	1191	20.4	373	12	US-10-408-967-8	Sequence 8, Appli
18	1191	20.4	373	14	US-10-060-036-72	Sequence 72, Appl
19	1183	20.2	373	16	US-10-466-258-4	Sequence 4, Appli
20	925	15.8	199	9	US-09-893-348-21	Sequence 21, Appl
21	908	15.5	199	9	US-09-893-348-25	Sequence 25, Appl
22	908	15.5	199	12	US-10-660-946-1	Sequence 1, Appli
23	901	15.4	199	12	US-10-408-967-9	Sequence 9, Appli
24	899	15.4	199	12	US-09-978-360A-467	Sequence 467, App
25	888	15.2	289	9	US-09-789-386-4	Sequence 4, Appli
26	868	14.8	199	16	US-10-466-258-11	Sequence 11, Appl
27	801	13.7	777	14	US-10-205-219-93	Sequence 93, Appl
28	792	13.5	780	12	US-10-267-502-432	Sequence 432, App
29	789.5	13.5	776	12	US-10-660-946-5	Sequence 5, Appli
30	789.5	13.5	776	12	US-10-267-502-430	Sequence 430, App
31	716.5	12.3	356	12	US-10-660-946-6	Sequence 6, Appli
32	706.5	12.1	593	15	US-10-108-260A-2892	Sequence 2892, Ap
33	685	11.7	208	12	US-10-660-946-7	Sequence 7, Appli
34	671	11.5	267	12	US-10-660-946-8	Sequence 8, Appli
35	671	11.5	267	14	US-10-205-194-127	Sequence 127, App
36	630	10.8	266	12	US-10-276-774-2330	Sequence 2330, Ap
37	625.5	10.7	236	9	US-09-729-674-20	Sequence 20, Appl
38	625.5	10.7	236	9	US-09-765-205-26	Sequence 26, Appl
39	625.5	10.7	236	12	US-10-408-967-2	Sequence 2, Appli
40	625.5	10.7	269	14	US-10-106-698-6222	Sequence 6222, Ap
41	539.5	9.2	168	10	US-09-809-391-563	Sequence 563, App
42	539.5	9.2	168	10	US-09-882-171-563	Sequence 563, App
43	539.5	9.2	168	12	US-10-164-861-563	Sequence 563, App
44	520	8.9	222	12	US-10-267-502-428	Sequence 428, App
45	519	8.9	241	12	US-10-660-946-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
 US-09-893-348-18  
 ; Sequence 18, Application US/09893348  
 ; Patent No. US20020072493A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EISENBACH-SCHWARTZ, Michal

```
; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-893-348-18
```

```
Query Match          100.0%; Score 5848; DB 9; Length 1163;
Best Local Similarity 100.0%; Pred. No. 6.1e-287;
Matches 1163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MEDIDQSSILVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60
          |||
Db      1 MEDIDQSSILVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60

Qy      61 PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
          |||
Db      61 PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120

Qy      121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALP 180
          |||
Db      121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALP 180

Qy      181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSVFKEHGYL 240
          |||
Db      181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSVFKEHGYL 240

Qy      241 GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSFSGSPKGES 300
          |||
Db      241 GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSFSGSPKGES 300

Qy      301 AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA 360
          |||
Db      301 AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA 360

Qy      361 PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEGR 420
          |||
Db      361 PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEGR 420
```

Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLSLTVFSIVSVTAYIALALLS	1020
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Db	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Qy	1141	NKSVKDAMAKIQAKIPGLKRRAD	1163
Db	1141	NKSVKDAMAKIQAKIPGLKRRAD	1163

RESULT 2

US-10-267-502-431

; Sequence 431, Application US/10267502

; Publication No. US20040071700A1

```
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 431
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-267-502-431
```

```
Query Match          90.7%; Score 5307; DB 12; Length 1163;
Best Local Similarity 91.1%; Pred. No. 1.3e-259;
Matches 1066; Conservative 34; Mismatches 56; Indels 14; Gaps 7;
```

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
          |||
Db      1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEEEVLERK 59

Qy     61 PAAGLSAAAVPPAAAAPLLDfSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
          |||
Db     60 PAAGLSAVPVFP-AAAPLLDfSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP 118

Qy    121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETlFALP 180
          |||
Db    119 PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAAPPSTPAAPKRRGSGSVDETlFALP 176

Qy    181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGyl 240
          |||
Db    177 AASEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFETAASLPSLSPLSTVSFKEHGyl 236

Qy    241 GNLSAVSSSEGTIEETlNEASKELPERATNPFVNRDLAEFSELEYSEMGSsfKGSfPKGES 300
          |||
Db    237 GNLSAVASTEGTIEETlNEASRELPERATNPFVNRRESAEFSVLEYSEMGSsfNGSPKGES 296

Qy    301 AILVENTKEEVIVRSKDKEDLVCSAALHSPQESP-----VGKEDRVVSPEKTMDIFNEMQ 355
          |||
Db    297 AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATlTKVVKEDGVMSPEKTMDIFNEMK 356

Qy    356 MSVVAPVREEYADfKPFQAWEVKDTYEGSRDVLAAANVESKVDRCLEDsLEQKSLGK 415
          |||
Db    357 MSVVAPVREEYADfKPFQAWEVKDTYEGSRDVLAAANMESKVDKKCFEDsLEQKSHGK 416

Qy    416 DSEGRNEDASFPSTPEPVKDSSRAYITCASfTSATESTANTfPllEDHTSENKTDEKKI 475
          |||
Db    417 DSESRNENASFPSTPELVKDGSRAYITCDSfTSATESTAANIFPVLEDHTSENKTDEKKI 476

Qy    476 EERKAQIIITEKTSPKTSNPFLVAVQDSEADYVTtDTLSKVTEAAVSNMPEGLTPDLVQEA 535
          |||
Db    477 EERKAQIIITEKTSPKTSNPFLVAIHdSEADYVTtDNLSKVTEAVVATMPEGLTPDLVQEA 536

Qy    536 CESELNEATGTKIAYETKVdlVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEA 595
          |||
```





; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 60/236,378  
; PRIOR FILING DATE: 2000-09-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-140-6

Query Match 75.3%; Score 4403.5; DB 9; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 5.9e-214;  
Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |||:||||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAAPLLDFSSDVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| |||:| | ||||| |||||: || |
Db      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy      116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||| ||||| ||||| ||||| || || |||||
Db      119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPPSTPAAPKR 178

Qy      167 RG-SGSVDETLEFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
      || ||||| ||||| |||| ||| |||||:|:|||||
Db      179 RGSSGSVDETLEFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy      226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
      |||| |||| ||||| | :|||:| :|||:| |:| :||| |||||
Db      238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy      286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
      ||||| ||| |||:| | :||:|:|:|:| | : ||: || | |||
Db      298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED 357

Qy      340 RVVSPEKTMDFNEMQMSVVPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV 395
      ||| || | ||| :|:| ||:|||||||: ||||: | | |:| :|:
Db      358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 416

Qy      396 ESKVDRKCLEDSELEQSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454
      ||||:| |||| : |||| |:| ||||| :|| | ||||| | :|||
Db      417 ESKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA 476

Qy      455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513
      | |||| | ||||| |||||:||||:| || ||||| |||| ||||| |:
Db      477 TNIFPLLGDPTSENKTDEKKIEEKKQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy      514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL 573
      |||| |:||||| ||||| |||||:||||| :||| |||
Db      537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL 596

Qy      574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPVSYDSIKLEPEN 633
      |||||:||||| ||||| :|||||:| |||| |:|:| |||
```



```
; PRIOR APPLICATION NUMBER: 60/175,707
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-599A-6
```

Qy	1	MEDIDQSSSLVSSSDSPPRPPPAFKYQFVTEPEDEEDEEEEDDEEDLEEELEVLERK	60
Db	1	MEDLDQSPLVSSS--DSPPRPQPAFKYQFVREPEDEE-EEEEEEEDDEDEDLEEELEVLERK	58
Qy	61	PAAGLSAAAVP--PAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVP	118
Qy	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS	237
Qy	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPVNRDLAEFSELEY	285
Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESVPG-----KED	339
Db	298	SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMIDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAA----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA	476
Qy	455	ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS	513
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNL	536
Qy	514	KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
Db	537	KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSIKLEPEN	633
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN	655

Qy 634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDF 692  
 Db 656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDF 715

Qy 693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS 752  
 Db 716 SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPEVPQKQDETVMVLKESLTETS 775

Qy 753 -ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF 809  
 Db 776 FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL 835

Qy 810 NTAIYSNDDLSSKEDIKESSETFSDSSPIEIIDEFPTFVSAKDDS-PKLAKEYTDLEVS 868  
 Db 836 STAVYSNDDLFIKSEAQIRETETFSDDSSPIEIIDEFPTLISSKTDSESKLAREYTDLEVS 895

Qy 869 DKSEIANIQSGADSLPCLLEPCDLSEFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV 926  
 Db 896 HKSEIANAPDGAGSLPCTELPHDLSEKNIQPKVEEKISFSDDFSKNGSATSQVLLLPDPV 955

Qy 927 SALEPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK 986  
 Db 956 SALATQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK 1015

Qy 987 KTGCVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA 1046  
 Db 1016 KTGCVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA 1075

Qy 1047 YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1106  
 Db 1076 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1135

Qy 1107 NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 1136 NGLTLLILALISLFSVPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 1192

# RESULT 5

US-10-060-036-71

; Sequence 71, Application US/10060036

; Publication No. US20030073144A1

; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Persing, David H.

; APPLICANT: Hepler, William T.

; APPLICANT: Jiang, Yuqiu

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER

; FILE REFERENCE: 210121.566

; CURRENT APPLICATION NUMBER: US/10/060,036

; CURRENT FILING DATE: 2002-01-30

; NUMBER OF SEQ ID NOS: 4560

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 71

; LENGTH: 1192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-060-036-71

Query Match 75.3%; Score 4403.5; DB 14; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 5.9e-214;  
Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEDEDEDEDEDEDEDEDEDEDELEELVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDELEELVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA--P 115
      ||||| || ||| |||:|:|:| ||||| ||||| |||||: ||: |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVP 118

Qy    116 APSLPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSVDETLFALPAASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
      || ||||| ||||| ||||| ||| ||| |||||:|:| |||||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy    226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
      ||||| ||||| ||||| | :|||:| :|||:| |:| :||| |||||
Db    238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy    286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
      ||||| ||| |||:|:| | :||:|:|:|:|: || : ||: || | |||
Db    298 SEMGSSFVSFPAESAVIVANPREELIVKNKDDEEKLVSNNILHNQQELPTALTCLKVKED 357

Qy    340 RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWVVDYEGSRDVLAA----RANV 395
      ||| || | ||| :||:|:| |||||: ||||: | | |:||| :|:
Db    358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 416

Qy    396 ESKVDRKCLEDSELEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454
      |||||:| ||||| : |||| |:| ||||| :|| | ||||| | :|||
Db    417 ESKVDKKCFADSLEQTNHEKDSSESSNDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA 476

Qy    455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513
      | |||| | ||||| |||||:|:|:| || | ||||| |||| | ||||| |:
Db    477 TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy    514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTGIAYETKVDLVQTSEAIQESLYPTAQL 573
      |||| |:| ||||| ||||| |||||:|:| ||||| :||| |||
Db    537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTGIAYETKMDLVQTSEVMQESLYPAAQL 596

Qy    574 CPSFEEAEATPSVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSIKLEPEN 633
      |||||:| ||||| ||||| :|||:|:| ||||| |:|:| |||
Db    597 CPSFEESEATPSVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN 655

Qy    634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF 692
      |||||:|:|:| : | || ||||: ||:| ||||| ||||| ||:| |||
Db    656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDF 715
```



Query Match 75.2%; Score 4398.5; DB 9; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 1e-213;  
Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPEDEE-EEEEEEDEDEDLEEEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| |||:| | ||||| |||||: || : |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| ||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
      || ||||| ||||| ||||| ||| |||||:|:| |||||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy    226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
      ||||| ||||| ||||| | :||:| :|||: |:| :||| |||||
Db    238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy    286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
      ||||| ||| |||:| | :||:|:|:|:| | : ||: || | |||
Db    298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQOELPTALTCLKVKED 357

Qy    340 RVVSPEKTMDIFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA-----RANV 395
      ||| || | ||| ::| ||:|||||||: ||||: | | |:||| :|:
Db    358 EVVSSEKAKDSFNEKRAVEAPMREEYADFKPFERWVEVKDSKEDS-DMLAAGGKIESNL 416

Qy    396 ESKVDRKCLEDSELEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454
      ||||:| |||| : |||| |:| ||||| :|| ||||| | :|||
Db    417 ESKVDKCCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA 476

Qy    455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513
      | |||| | ||||| |||||:|||:| || ||||| ||||| ||||| |:
Db    477 TNIFPLLGDPTSSENKTDEKKIEEKKAIIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy    514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSIAIQESLYPTAQL 573
      |||| |:||||| ||||| ||||| |||||:||||| :||||| |||
Db    537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSVQMESLYPAAQL 596

Qy    574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVPLEAPPPVSYDSIKLEPEN 633
      |||||:||||| ||||| :|||||:| |||| | |:| ||| |||
Db    597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPEN 655

Qy    634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSDF 692
      |||||:|:| | : | || ||||: ||:||||| ||||| |||:|
Db    656 PPPYEEAMSVSLKKVSGIKKEIKPENINAALQETEAPYISIACDLIKETKLSAEPAPDF 715

Qy    693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVS 752
      |:|:|:| | : |:|:|:|||||:||||| ||| |:| |:|:|
Db    716 SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLFSDDSIPEVPQKQDETVMVKESLTETS 775
```





; ORGANISM: Homo sapiens  
US-09-893-348-23

Query Match 75.2%; Score 4398.5; DB 9; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 1e-213;  
Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

```
Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEELEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || ||| |||:|:|:| ||||| |||||: || : |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDVFPPAPRGPLPAAPPVAPERQPSWDPSVSVSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| || ||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RG-SGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
      || ||||| ||||| ||||| ||| |||||:|:| |||||
Db    179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy    226 LSPLSTVSFKEHGYLGNLSAVSSSEG TIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
      ||||| ||||| ||||| | :|||:| :|||: |:| :||| |||||
Db    238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy    286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
      ||||| ||| |||:| | :||:|:|:|:| | : ||: || | |||
Db    298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDDEEEKLVSNNILHNQQELPTALTCLKVKED 357

Qy    340 RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV 395
      ||| || | ||| :||| ||:|||||||: ||||: || | ||| :|:
Db    358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL 416

Qy    396 ESKVDRKCLED SLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454
      ||||:| |||| : |||| |:| ||||| :|| ||||| | :|||
Db    417 ESKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA 476

Qy    455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513
      | |||| | ||||| |||||:||||: ||| || ||||| ||| ||||| |:
Db    477 TNIFPLLGDPTSENKTDEKKIEEKKQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy    514 KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL 573
      |||| |:||||| ||||| ||||| |||||:||||| :||| |||
Db    537 KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL 596

Qy    574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN 633
      |||||:||||| ||||| :|||||:||| |||| |:|:| |||
Db    597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPS SSPLEA-SSVNYESIKHEPEN 655

Qy    634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDF 692
      |||||:|:|:| : | || ||||: ||:||||| ||||| ||:|
Db    656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDF 715

Qy    693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVS 752
      |:|:|:| |:|:|:| |||||:||||| |||||:| |:| ||:| |||||
```

Db	716	SDYSEMAKVEQVPVDPHSELVEDSSPDSEPVDLFSDDSIDPVPQKQDETVMVLVKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
Db	776	FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPFVSAKDDS-PKLAKEYTDLEVS	868
Db	836	STAVYSNDDLFIKAEQIRETETFSFSDSSPIEIIIDEFPTLISSKTDSEKSLAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLLPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
Db	896	HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDPV	955
Qy	927	SALEPQTEMGSIVKSKSLTKEAEKKLPSTTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	986
Db	956	SALATQAEIESIVKPKVLVKEAEKKLPSTTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK	1015
Qy	987	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1046
Db	1016	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1075
Qy	1047	YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF	1106
Db	1076	YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF	1135
Qy	1107	NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD	1163
Db	1136	NGLTLLILALISLFSVPVIYERHQVQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192

### RESULT 8

; Sequence 429, Application US/10267502

: Publication No. US20040071700A1

; GENERAL INFORMATION:

: APPLICANT: Kim, Jaeseob

; APPLICANT: Galant, Ron

: TITLE OF INVENTION: Obesity Linked Genes

; FILE REFERENCE: LSD-07416

: CURRENT APPLICATION NUMBER: US/10/267,502

: CURRENT FILING DATE: 2003-01-27

; NUMBER OF SEQ ID NOS: 439

```

; SOFTWARE: PatentIn version 3.2

```

; SEO ID NO 429

; LENGTH: 1192

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-267-502-429

Query Match 75.2%; Score 4398.5; DB 12; Length 1192;

Best Local Similarity 75.9%; Pred. No. 1e-213;

Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

Qy 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60

1 MEDLDQSPVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58

Qy 61 PAAGLSAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
 ||||| || ||| ||:| :| ||||| ||||| ||||| : || : |  
 Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSSTVP 118

Qy 116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
 ||| |||| ||||| ||||| ||||| || || ||||| |||||  
 Db 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy 167 RG-SGSVDETLFALPAASEPVISSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225  
 || ||||| ||||| ||||| ||| ||| ||||| :|: ||||| |||||  
 Db 179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy 226 LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285  
 ||||| ||||| ||||| | :||:| :||:| :| :||| |||||  
 Db 238 LSPLSAAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY 297

Qy 286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339  
 ||||| ||| |||:| | :||:| :||:| :| : ||: || | |||  
 Db 298 SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED 357

Qy 340 RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDLAA----RANV 395  
 ||| || | ||| :|:| ||:||||||| : ||||: | | |:| | :|:  
 Db 358 EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVVEVKDSKEDS-DMLAAGGKIESNL 416

Qy 396 ESKVDRKCLEDLSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT 454  
 ||||:| |||| : |||| |:| ||||| :|| ||||| | :|||  
 Db 417 ESKVDKKCFADSLEQTNHEKDSSESSNDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA 476

Qy 455 ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS 513  
 | |||| | ||||| |||||:||||:| || ||||| ||||| ||||| |:  
 Db 477 TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT 536

Qy 514 KVTEAAVSNMPEGLTPDLVQFACESELNEATGTKIAYETKVDLVQTEAIIQESLYPTAQL 573  
 |||| |:||||| ||||| ||||| |||||:||||| :||||| |||  
 Db 537 KVTEEVVANMPEGLTPDLVQFACESELNEVTGTKIAYETKMDLVQTEVMQESLYPAAQL 596

Qy 574 CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLAAPPVSYDSIKLEPEN 633  
 |||||:||||| ||||| : |||||:| ||||| |:| ||| ||||  
 Db 597 CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN 655

Qy 634 PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDF 692  
 |||||:|:| | : | || ||||: |||:||||| ||||| ||:|  
 Db 656 PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISACDLIKETKLSAEPAPDF 715

Qy 693 SNYSEIAKFEKSVPEHAELVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVS 752  
 |:|:|:| |: ||:|:|||||:||||| ||| |:| |||:||||| |  
 Db 716 SDYSEMAKVEQVPVDPHSELVEDSSPDSEPVDFLSDDSIPDVPQKQDETVMVLKESLTETS 775

Qy 753 -ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF 809  
 |:| :|:| :|:| | | |||||:| :|| :|:| ||:| ||||  
 Db 776 FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL 835

Qy 810 NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPTFVSAKDDS-PKLAKEYTDLEVS 868  
 :|:| |||| | || :|:| ||||| ||||| :|:| || |||||  
 Db 836 STAVYSNDDLFIKESQIRETETFSDDSSPIEIIIDEFPTLISSKTDSESKLAREYTDLEVS 895

Db 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy	167	RG-SGSVDETLFALPAASEPVIIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS	225
		:	
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS	237
Qy	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
		:    :    :    :    :	
Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMSSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
		:    :    :    :    :    :	
Db	298	SEMSSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA----RANV	395
		:    :    :    :    :    :	
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDRCLEDLSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
		:          :          :          :	
Db	417	ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA	476
Qy	455	ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS	513
		:    :    :                      :	
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT	536
Qy	514	KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
		:    :    :    :    :    :    :    :	
Db	537	KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN	633
		:    :    :    :    :    :    :    :	
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN	655
Qy	634	PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF	692
		:    :    :    :    :    :    :    :	
Db	656	PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSEAEPAPDF	715
Qy	693	SNYSEIAKFEKSVPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVS	752
		:    :    :    :    :    :    :    :    :	
Db	716	SDYSEMAKVEQVPDHSSELVEDSSPDSEPVDLFSDDSIPEVPQKQDETVMVLKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
		:    :    :    :    :    :    :    :    :	
Db	776	FESMIEYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDEFPTFVSAKDDS-PKLAKEYTDLEVS	868
		:    :    :    :    :    :    :    :    :	
Db	836	STAVYSNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSEFSKLAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLLEPCDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
		:    :    :    :	
Db	896	HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDPV	955
Qy	927	SALEPQTEMGSIKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	986
		:    :    :    :    :    :    :    :	
Db	956	SALATQAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK	1015

```

Qy      987 KTG VVFGASL FLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1046
          |||
Db      1016 KTG VVFGASL FLLLSLT VFSIVSVTAYIALALLSVTISFRIYKGV IQAIQKSDEGHPFRA 1075

Qy      1047 YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1106
          |||
Db      1076 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF 1135

Qy      1107 NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163
          |||:||||| |||:|||||:|||||:
Db      1136 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

```

RESULT 10

US-10-466-258-9

```

; Sequence 9, Application US/10466258
; Publication No. US20040132096A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P80966 GCW
; CURRENT APPLICATION NUMBER: US/10/466,258
; CURRENT FILING DATE: 2003-07-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-258-9

```

```

Query Match          75.2%; Score 4398.5; DB 16; Length 1192;
Best Local Similarity 75.9%; Pred. No. 1e-213;
Matches 908; Conservative 104; Mismatches 146; Indels 39; Gaps 20;

```

```

Qy      1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
          |||:|||| |||| |||| |||| |||| ||||:||:|||||
Db      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAPLDDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
          ||||| || ||| |||:| :| ||||| ||||| |||||: || :|
Db      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVVSSTVP 118

Qy      116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
          ||| |||| ||||| ||||| |||| || || |||||
Db      119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy      167 RG-SGSVDETLFALPAASEPVI PSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225
          || ||||| ||||| |||| ||| |||||:|:|||||
Db      179 RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237

Qy      226 LSPLSTVSFKEHGYLGNL SAVSSSEG TIEETLNEASKELPERATNPFVNRDLAEFSELEY 285
          ||||| ||||| ||||| | :||:| :|||:|:| :||| |||||
Db      238 LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRLTEFSELEY 297

Qy      286 SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED 339
          ||||| ||| |||:| | :||:|:|:|:| | : ||: || | |||

```

Db	298	SEMGSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWVEKDTYEGSRDLAA----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSSESNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIA	476
Qy	455	ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS	513
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT	536
Qy	514	KVTEAAVSNMPEGLTPDLVQFACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
Db	537	KVTEEVVANMPEGLTPDLVQFACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN	633
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQSSSPLEA-SSVNYESIKHEPEN	655
Qy	634	PPPYEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF	692
Db	656	PPPYEAMSVSLLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDF	715
Qy	693	SNYSEIAKFESVPEHAELVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVS	752
Db	716	SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDLSDDSIPDVPQKQDETVMVLKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
Db	776	FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIDEFPTFVSAKDDS-PKLAKEYTDLEVS	868
Db	836	STAVYSNDDLFIKSEAQIRETETFSDDSSPIEIDEFPTLISSKTDIFSGLAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLLPCLDLSFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
Db	896	HKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSNGSATSQVLLLPDPV	955
Qy	927	SALEPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	986
Db	956	SALATQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK	1015
Qy	987	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1046
Db	1016	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1075
Qy	1047	YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF	1106
Db	1076	YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF	1135
Qy	1107	NGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD	1163
Db	1136	NGLTLLILALISLFSVPVIYERHQVQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	1192



US-10-408-967-7

Query Match 75.1%; Score 4389.5; DB 12; Length 1192;  
Best Local Similarity 75.8%; Pred. No. 3e-213;  
Matches 907; Conservative 104; Mismatches 147; Indels 39; Gaps 20;

Qy	1	MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEDEEDEDLEELEVLERK	60
Db	1	MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK	58
Qy	61	PAAGLSAAAVP--PAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVPSSTVP	118
Qy	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSVDETLFALPAASEPVI PSSAEKIMDLMEQPNTVSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSAEN-MDLKEQPNTISAGQEDFPSVLLETAASLPS	237
Qy	226	LSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMGSSFKGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
Db	298	SEMGSSFSVSPKAESAVIVANPREEIIIVKNKDEEEKLVSNNILHNQQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWVEVKDITYEGSRDVLAA----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDRKCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSRAYITCASFT-SATESTT	454
Db	417	ESKVDDKKCFADSLEQTNHEKDESSNDDTSFSPSTPEGIKDRPGAYITCAPFNPAATESIA	476



```
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 164
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Foocen-m2 reticulon
US-10-205-194-164
```

```
Query Match          25.9%; Score 1513; DB 14; Length 379;
Best Local Similarity 32.6%; Pred. No. 1.4e-68;
Matches 379; Conservative 0; Mismatches 0; Indels 784; Gaps 1;
```

```
Qy      1 MEDIDQSSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60
        |
Db      1 MEDIDQSSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60

Qy     61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
        |
Db     61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120

Qy    121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP 180
        |
Db    121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALP 180

Qy    181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL 240
        |
Db    181 AASEPVIPSSA----- 191

Qy    241 GNLSAVSSSEGTIEETLNEASKELPERATNPVNRDLAEFSELEYSEMGSSEFKGSPKGES 300
        |
Db    192 ----- 191

Qy    301 AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA 360
        |
Db    192 ----- 191

Qy    361 PVREEYADFKPFQAWVEVKDTYPEGSRDVLAARANVESKVDRKCLEDSELEQKSLGKDSEGR 420
        |
Db    192 ----- 191

Qy    421 NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA 480
        |
Db    192 ----- 191

Qy    481 QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL 540
```

Db	192	-----	191
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	192	-----	191
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	192	-----	191
Qy	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	192	-----	191
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	192	-----	191
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	192	-----	191
Qy	841	IIDEFPPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	192	-----	191
Qy	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	192	-----	191
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	192	-----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	236
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	237	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	296
Qy	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	1140
Db	297	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGIA	356
Qy	1141	NKSVKDAMAKIQAKIPGLKRRKAD	1163
Db	357	NKSVKDAMAKIQAKIPGLKRRKAD	379

RESULT 13

US-09-893-348-20

; Sequence 20, Application US/09893348

; Patent No. US20020072493A1

; GENERAL INFORMATION:

; APPLICANT: EISENBACH-SCHWARTZ, Michal

; APPLICANT: COHEN, Irun R.

; APPLICANT: BESERMAN, Pierre

; APPLICANT: MOSONEGO, Alon

; APPLICANT: MOALEM, Gila



Qy	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	173	-----	172
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	173	-----	172
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	173	-----	172
Qy	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	173	-----	172
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	173	-----	172
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	173	-----	172
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	173	-----	172
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	173	-----	172
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	173	-----VVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	217
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	218	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	277
Qy	1081	FLVDDLVDLSLKFVILMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Db	278	FLVDDLVDLSLKFVILMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	337
Qy	1141	NKSVKDAMAKIQAKIPGLKRKAD	1163
Db	338	NKSVKDAMAKIQAKIPGLKRKAD	360

RESULT 14

US-09-789-386-6

; Sequence 6, Application US/09789386

; Patent No. US20020010324A1

; GENERAL INFORMATION:

; APPLICANT: MICHALOVICH, DAVID

; APPLICANT: PRINJHA, RABINDER KUMAR

; TITLE OF INVENTION: NOVEL COMPOUNDS

```
; FILE REFERENCE: GP-30165-C1
; CURRENT APPLICATION NUMBER: US/09/789,386
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: U.K. 9916898.1
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: U.K. 9816024.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: US 09/359,208
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-789-386-6
```

```
Query Match          20.4%; Score 1191; DB 9; Length 373;
Best Local Similarity 27.8%; Pred. No. 2.5e-52;
Matches 327; Conservative 12; Mismatches 20; Indels 818; Gaps 7;
```

```
Qy      1 MEDIDQSSSLVSSSTDSPRRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
      |||:||| ||||| ||||| ||||| ||||| ||||| |||||:|||:||| ||||| |||||
Db      1 MEDLDQSPLVSSS-DSPRRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEELEVLERK 58

Qy      61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || ||| |||:| | :| ||||| ||||| |||||:|||: || : |
Db      59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy      116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEVWTPAPAPAAPSTPAAPKR 178

Qy      167 RGSGSVDETLFALPAASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL 226
      |||
Db      179 RGSS----- 182

Qy      227 SPLSTVSFKEHGYLGNL SAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYS 286
Db      183 ----- 182

Qy      287 EMGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESVPGKEDRVVSPEK 346
Db      183 ----- 182

Qy      347 TMDIFNEMQMSV VAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLED 406
Db      183 ----- 182

Qy      407 SLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTS 466
Db      183 ----- 182

Qy      467 ENKTDEKKIEERKAQIIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEG 526
Db      183 ----- 182
```

Qy 527 LTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSPFEEAEATPSP 586  
 Db 183 ----- 182  
 Qy 587 VLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPVSYDSIKLEPENPPPYEEAMNVALK 646  
 Db 183 ----- 182  
 Qy 647 ALGTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSV 706  
 Db 183 ----- 182  
 Qy 707 EHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSAS 766  
 Db 183 ----- 182  
 Qy 767 PQELGKPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDK 826  
 Db 183 ----- 182  
 Qy 827 IKESETFSDSSPIEIIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCL 886  
 Db 183 ----- 182  
 Qy 887 ELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTK 946  
 Db 183 -----GSV----- 185  
 Qy 947 EAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS 1006  
 Db 186 -----VVDDLLYWRDIKKTGVVFGASLFLLLSLTVFS 216  
 Qy 1007 IVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA 1066  
 Db 217 IVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA 276  
 Qy 1067 LGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALENGLTLLILALISLFSIPVIY 1126  
 Db 277 LGHVNTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALENGLTLLILALISLFSVPVIY 336  
 Qy 1127 ERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 337 ERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 373

RESULT 15

US-09-765-205-6

; Sequence 6, Application US/09765205

; Patent No. US20020034800A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Li

; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES

; FILE REFERENCE: 1458.004/200130.449

; CURRENT APPLICATION NUMBER: US/09/765,205

; CURRENT FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: US/09/212,440

; PRIOR FILING DATE: 1998-12-16



; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: human  
US-09-765-205-6

Query Match 20.4%; Score 1191; DB 9; Length 373;  
Best Local Similarity 27.8%; Pred. No. 2.5e-52;  
Matches 327; Conservative 12; Mismatches 20; Indels 818; Gaps 7;

```
Qy      1 MEDIDQSSLVSSSTDSPRRPPPAFKYQFVTEPEDEDEEEEEDEEDEDLEEEVLERK 60
      |||:|||| ||||| ||||| ||||| ||||| |||||:|:|:|||||
Db      1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK 58

Qy     61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115
      ||||| || || |||:|:|:| ||||| |||||: ||: |
Db     59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSTVP 118

Qy    116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166
      ||| |||| ||||| ||||| ||||| ||||| |||||
Db    119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKR 178

Qy    167 RGSGSVDETFLFALPAASEFVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSL 226
      |||
Db    179 RGSS----- 182

Qy    227 SPLSTVSFKEHGYLGNLSAVSSSEGTIEETLINEASKELPERATNPFVNRDLAEFSELEYS 286
Db    183 ----- 182

Qy    287 EMGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEK 346
Db    183 ----- 182

Qy    347 TMDIFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANVESKVDRKCLED 406
Db    183 ----- 182

Qy    407 SLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTS 466
Db    183 ----- 182

Qy    467 ENKTDEKKIEERKAQIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEG 526
Db    183 ----- 182

Qy    527 LTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSP 586
Db    183 ----- 182

Qy    587 VLPDIVMEAPLNSLLPSAGASVVQPSVSPLAPPVSYDSIKLEPENPPPYEEAMNVALK 646
Db    183 ----- 182

Qy    647 ALGTKEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVP 706
```

Db	183	-----	182
Qy	707	EHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSAS	766
Db	183	-----	182
Qy	767	PQELGKPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDK	826
Db	183	-----	182
Qy	827	IKESETFSDSSPIEIIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCL	886
Db	183	-----	182
Qy	887	ELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTK	946
		:	
Db	183	-----GSV-----	185
Qy	947	EAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS	1006
Db	186	-----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFS	216
Qy	1007	IVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA	1066
Db	217	IVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA	276
Qy	1067	LGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFETYVGALFNGLTLLILALISLFSIPVIY	1126
Db	277	LGHVNCTIKELRRLFLVDDLVDLSLKFAVLMWVFETYVGALFNGLTLLILALISLFSVPVIY	336
Qy	1127	ERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD	1163
Db	337	ERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE	373

Search completed: September 29, 2004, 18:48:11  
Job time : 131.84 secs

Result		Query				Description	
No.	Score	Match	Length	DB	ID		

1	5312.5	90.8	1162	11	Q8BGM9	Q8bgm9 mus musculu
2	5307	90.7	1163	11	Q8K3G8	Q8k3g8 mus musculu
3	4501.5	77.0	1046	11	Q8BGK7	Q8bgk7 mus musculu
4	3627.5	62.0	986	4	Q8IUA4	Q8iua4 homo sapien
5	3299.5	56.4	720	11	Q7TNB7	Q7tnb7 mus musculu
6	2926	50.0	639	11	Q8K290	Q8k290 mus musculu
7	2610	44.6	578	11	Q80W95	Q80w95 mus musculu
8	1416	24.2	375	11	Q8BHF5	Q8bhf5 mus musculu
9	1314.5	22.5	356	11	Q8BH78	Q8bh78 mus musculu
10	1304	22.3	357	11	Q8K3G7	Q8k3g7 mus musculu
11	1283.5	21.9	392	4	Q96B16	Q96b16 homo sapien
12	878	15.0	184	6	Q7YRW9	Q7yrw9 bos taurus
13	876	15.0	199	13	Q7T224	Q7t224 gallus gall
14	844	14.4	179	6	Q9GM33	Q9gm33 macaca fasc
15	792	13.5	780	11	Q8K4S4	Q8k4s4 mus musculu
16	788	13.5	780	11	Q8K0T0	Q8k0t0 mus musculu
17	762.5	13.0	760	13	Q90638	Q90638 gallus gall
18	737	12.6	214	13	Q7T222	Q7t222 carassius a
19	700	12.0	643	11	Q8CCU2	Q8ccu2 mus musculu
20	685	11.7	199	4	Q9BQ59	Q9bq59 homo sapien
21	671	11.5	267	11	Q63765	Q63765 rattus sp.
22	669	11.4	208	13	Q90637	Q90637 gallus gall
23	625.5	10.7	236	11	Q8VBU0	Q8vbu0 rattus norv
24	625.5	10.7	237	11	Q8C6D5	Q8c6d5 mus musculu
25	590.5	10.1	221	13	Q7ZUD6	Q7zud6 brachydanio
26	586	10.0	595	5	Q9VMV9	Q9vmv9 drosophila
27	532	9.1	224	5	Q9VMW1	Q9vmw1 drosophila
28	520	8.9	202	5	Q9VMW2	Q9vmw2 drosophila
29	520	8.9	222	5	Q9VMW4	Q9vmw4 drosophila
30	520	8.9	234	5	Q9VMW3	Q9vmw3 drosophila
31	517	8.8	2484	5	Q9U347	Q9u347 caenorhabdi
32	503.5	8.6	2607	5	Q23187	Q23187 caenorhabdi
33	377	6.4	2768	5	Q9VC00	Q9vc00 drosophila
34	347.5	5.9	10578	5	Q8ISF5	Q8isf5 caenorhabdi
35	344	5.9	107	13	Q7T223	Q7t223 carassius a
36	343.5	5.9	18519	5	Q8ISF6	Q8isf6 caenorhabdi
37	343.5	5.9	18534	5	Q8ISF7	Q8isf7 caenorhabdi
38	342.5	5.9	1417	3	Q871Y7	Q871y7 neurospora
39	335	5.7	5412	5	Q9W596	Q9w596 drosophila
40	331	5.7	4900	5	Q9N541	Q9n541 caenorhabdi
41	328.5	5.6	5327	5	Q76891	Q76891 drosophila
42	322	5.5	7962	4	Q10465	Q10465 homo sapien
43	320	5.5	222	5	Q23188	Q23188 caenorhabdi
44	317.5	5.4	17352	5	Q95YM2	Q95ym2 procambarus
45	313.5	5.4	1444	5	Q9VTN2	Q9vtn2 drosophila

# ALIGNMENTS

## RESULT 1

Q8BGM9

ID Q8BGM9 PRELIMINARY; PRT; 1162 AA.

AC Q8BGM9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE RTN4.  
 GN RTN4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvcJ7, and 129SvcJ7;  
 RA Oertle T., van der Putten H., Schwab M.E.;  
 RT "Genomic Structure and Functional Characterization of the Promoter  
 RT Structures of Human and Mouse Nogo/Rtn-4.";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvcJ7, and 129SvcJ7;  
 RA Oertle T., Schwab M.E.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvcJ7;  
 RA Van der Putten H.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129SvcJ7;  
 RA Van der Putten H., Mir A.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY102284; AAM73506.1; -.  
 DR EMBL; AY102286; AAM73511.1; -.  
 DR MGD; MGI:1915835; Rtn4.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.  
 DR GO; GO:0007399; P:neurogenesis; IDA.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 SQ SEQUENCE 1162 AA; 126613 MW; 855697FBEE11781F CRC64;

Query Match 90.8%; Score 5312.5; DB 11; Length 1162;  
 Best Local Similarity 91.0%; Pred. No. 9.1e-259;  
 Matches 1064; Conservative 37; Mismatches 55; Indels 13; Gaps 6;

Qy	1	MEDIDQSSIVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Db	1	MEDIDQSSIVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEELEVLERK	59
Qy	61	PAAGLSAAVPPAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP	180
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP	176
Qy	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	177	AASEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFETAASLPSLSPLSTVSFKEHGYL	236

Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
		: :     :     :	
Db	237	GNLSAVASTEGTIEETLNEASRELPERATNPFVNRESAEFSVLEYSEMGSSFNGSPKGES	296
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESP-----VGKEDRVVSPEKTMDFNEMQ	355
		:     :            :     :	
Db	297	AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDFNEMK	356
Qy	356	MSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDLSLEQKSLGK	415
		:     :     :	
Db	357	MSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANMESKVDDKKCFEDSLEQKGHGK	416
Qy	416	DSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKI	475
		:                  :          :	
Db	417	DSESRNENASFPRTPELVKDGSRAYITCDSFSSATESTAANIFPVLEDHTSENKTDEKKI	476
Qy	476	EERKAQIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEA	535
		:             :	
Db	477	EERKAQIITEKTS PKTSNPFLVAIH DSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEA	536
Qy	536	CESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEA	595
		:	
Db	537	CESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA	596
Qy	596	PLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIK	655
		:     :	
Db	597	PLNSLLPSTGASVAQPSASPLEVSPVSYDGIKLEPENPPPYEEAMSVALKTSDSKEEIK	656
Qy	656	EPESFNAAVQETEAPYIS IACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDS	715
		:     :     :	
Db	657	EPESFNAAAQAEAPYIS IACDLIKETKLSTEPSPEFSNYSEIAKFEKSVPDHCELVD DS	716
Qy	716	SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESL TEVSETVAQHK-EERLSASPQELGKPY	774
		:         :     :	
Db	717	SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESL TEVSETVTQHKHKERLSASPQEVGKPY	776
Qy	775	LESFQPNLHSTKDAASNDIPTLT KKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFS	834
		:           :	
Db	777	LESFQPNLHITKDAASNEIPTLT KKETISLQMEEFNTAIYSNDDLSSKEDKMKESSETFS	836
Qy	835	DSSPIEIIDFPTFVSAKDDSPKLAK EYTDLEVSDKSEIANIQSGADSLPCLELPCDL SF	894
		:     :     :     :	
Db	837	DSSPIEIIDFPTFVSAKDDSP---KEYTDLEVSNKSEIANVQSGANSLPCSELPCDL SF	893
Qy	895	KNIYPKDEVHVSDEFSEN RSSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPS	954
		:      :           :           :	
Db	894	KNTYPKDEAHVSDEF SKSRSSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEKKLPS	953
Qy	955	DTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYI	1014
		:     :	
Db	954	DTEKEDRSLTAVLSAELNKT SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYI	1013
Qy	1015	ALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTI	1074
Db	1014	ALALLSVTISFRIYKGV IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTI	1073

QY 1075 KELRRFLVDDLVDLSLKFAVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQVQID 1134  
 |||  
 Db 1074 KELRRFLVDDLVDLSLKFAVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQAQID 1133  
 QY 1135 HYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 |||  
 Db 1134 HYLGLANKSVKDAMAKIQAKIPGLKRKAE 1162

RESULT 2

Q8K3G8

ID Q8K3G8 PRELIMINARY; PRT; 1163 AA.  
 AC Q8K3G8;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Nogo-A.  
 GN RTN4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Jin W., Long M., Li R., Ju G.;  
 RT "Cloning and expression of the mouse Nogo-A protein.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY114152; AAM77068.1; -.  
 DR MGD; MGI:1915835; Rtn4.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.  
 DR GO; GO:0007399; P:neurogenesis; IDA.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 SQ SEQUENCE 1163 AA; 126691 MW; 6B5F362799417EA4 CRC64;

Query Match 90.7%; Score 5307; DB 11; Length 1163;  
 Best Local Similarity 91.1%; Pred. No. 1.7e-258;  
 Matches 1066; Conservative 34; Mismatches 56; Indels 14; Gaps 7;

QY 1 MEDIDQSSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60  
 |||  
 Db 1 MEDIDQSSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEDEDEDLEEELEVLERK 59  
 QY 61 PAAGLSAAAVPPAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120  
 |||  
 Db 60 PAAGLSAVPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP 118  
 QY 121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP 180  
 |||  
 Db 119 PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPPPSTPAAPKRRGSGSVDETLFALP 176  
 QY 181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL 240  
 |||  
 Db 177 AASEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFETAASLPSLSPLSTVSFKEHGYL 236

Qy 241 GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES 300  
 |||||:|:|||||||:|||||||: ||| ||||| |||||  
 Db 237 GNLSAVASTEGTIEETLNEASRELPERATNPFVNRSAEFSVLEYSEMGSSFNGSPKGES 296

Qy 301 AILVENTKEEVIVRSKDKEDLVCSAALHSPQESP-----VGKEDRVVSPEKTMDFNEMQ 355  
 |:|||||||:||||| | || |:|||||||:  
 Db 297 AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDFNEMK 356

Qy 356 MSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDLSLEQKSLGK 415  
 |||||:|||||:|| ||||| ||  
 Db 357 MSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAAARANMESKVDDKKCFEDSLEQKSHGK 416

Qy 416 DSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKI 475  
 ||| |||:||||| || ||||| || |||:|||||||  
 Db 417 DSESRNENASFPSTPELVKDGSRAYITCDSFTSATESTANIFPVLEDHTSENKTDEKKI 476

Qy 476 EERKAQIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEA 535  
 |||||: ||||| ||||| |: |||||  
 Db 477 EERKAQIITEKTS PKTSNPFLVAIHQDSEADYVTTDNLSKVTEAAVATMPEGLTPDLVQEA 536

Qy 536 CESELNEATGTKIAYETKVDLVQTS EAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEA 595  
 |||||:|||||:|||||  
 Db 537 CESELNEATGTKIAYETKVDLVQTS EAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA 596

Qy 596 PLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEAMNVALKALGTKEGIK 655  
 ||||| ||| ||| ||| |||||:|||| || ||  
 Db 597 PLNSLLPSTGASVAQPSASPLEVPSPVSYDGIKLEPENPPPYEAMSVALKTSDAKEEIK 656

Qy 656 EPESFNAAVQETEAPYIS IACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDS 715  
 ||||| || |||||:|||||: |||: ||  
 Db 657 EPESFNAAVQETEAPYIS IACDLIKETKLSTEPSPGFSNYSEIAKFEKSVPDHCELVDSDS 716

Qy 716 SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHK-EERLSASPQELGKPY 774  
 |||||:|||||: |||: |||||: |||||  
 Db 717 SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVTQHKHKERLSASPQEVGKPY 776

Qy 775 LESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFS 834  
 ||||| |||||:||||| |||||:|||||  
 Db 777 LESFQPNLHITKDAASNEIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKMKESSETFS 836

Qy 835 DSSPIEIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSF 894  
 |||||:|||||: |||||: |||||: |||||  
 Db 837 DSSPIEIDEFPTFVSAKDDSP---KEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSF 893

Qy 895 KNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPS 954  
 || |||| |||||: ||||| : ||||| |||: ||| |||||: |||||  
 Db 894 KNTYPKDEAHVSDEFSSKSRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEKKLPS 953

Qy 955 DTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAY 1013  
 |||||:|||||: ||||| |||||  
 Db 954 DTEKEDRSLTAVLSAELNKTSSVVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAY 1013

Qy 1014 IALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNST 1073  
 |||||:|||||: ||||| |||||  
 Db 1014 IALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNST 1073

Qy 1074 IKELRRLFLVDDLVDLSLKEAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQI 1133



```

Db      1074 IKELRRRLFLVDDLVDSLKFAVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQAI 1133
Qy      1134 DHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163
      :
Db      1134 DHYLGLANKSVKDAMAKIQAKIPGLKRKAE 1163

```

RESULT 3

Q8BGK7

```

ID   Q8BGK7          PRELIMINARY;          PRT;  1046 AA.
AC   Q8BGK7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7;
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SvcJ7;
RA   Van der Putten H., Mir A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY102280; AAM73502.1; -.
DR   EMBL; AY102286; AAM73507.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE  1046 AA;  114221 MW;  8CE2E2238ED51222 CRC64;

```

```

Query Match      77.0%;  Score 4501.5;  DB 11;  Length 1046;
Best Local Similarity  86.9%;  Pred. No. 4.8e-218;
Matches 910;  Conservative  40;  Mismatches  68;  Indels  29;  Gaps  4;

```

Qy	143	PPAGASPLAEPAPPPSTPA-----APKRRGSGSVDETLEFALPAA	182
Db	3	PPLAGGGQKGGAASEA WVPSLFVGVSGSTCTAAKSLVPIPARSSRLSAARNETLEFALPAA	62
Qy	183	SEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYLGN	242
Db	63	SEPVIPSSAEKIMDLKEQPGNTVSSGQEDFPSVLFEETAASLPSLSPLSTVSFKEHGYLGN	122
Qy	243	LSAVSSSEGTEIETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSEFKGSPKGESAI	302
Db	123	LSAVASTEGTEIETLNEASRELPERATNPFVNRRESAEFSVLEYSEMGSSEFNKSPKGESAM	182
Qy	303	LVENTKKEEVIVRSKDKEDLVCSAALHSPQESP-----VGKEDRVVSPEKTMDIFNEMQMS	357
Db	183	LVENTKKEEVIVRSKDKEDLVCSAALHNPQESPATLTQVVKEDGVMSPEKTMDIFNEMKMS	242
Qy	358	VVAPVREEYADFKPFQAWVEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEQKSLGKDS	417
Db	243	VVAPVREEYADFKPFQAWVEVKDTYEGSRDVLAAARANMESKVDRKCFEDSEQKSGHGKDS	302
Qy	418	EGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLEDHTSENKTDEKKIEE	477
Db	303	ESRNENASFPRTPELVKDGSRAYITCDSFSSATESTANIFPVLEDHTSENKTDEKKIEE	362
Qy	478	RKAQIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACE	537
Db	363	RKAQIITEKTS PKTSNPFLVAIHQDSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEACE	422
Qy	538	SELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSEEEAEATPSVLPDIVMEAPL	597
Db	423	SELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSEEEAEATPSVLPDIVMEAPL	482
Qy	598	NSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEP	657
Db	483	NSLLPSTGASVAQPSASPLEVPSPVSYDGIKLEPENPPPYEEAMSVALKTSKSDSKEEIKEP	542
Qy	658	ESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPHEAELVEDSSP	717
Db	543	ESFNAAAQEAAPYISIACDLIKETKLSTEPSPEFSNYSEIAKFEKSVPDHCELVDSSP	602
Qy	718	ESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHK-EERLSASPQELGKPYLE	776
Db	603	ESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVTQHKHKERLSASPQEVGKPYLE	662
Qy	777	SFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDS	836
Db	663	SFQPNLHITKDAASNEIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKMKESSETFSDS	722
Qy	837	SPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKN	896
Db	723	SPIEIIDEFPTFVSAKDDSP---KEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKN	779
Qy	897	IYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDST	956
Db	780	TYPKDEAHVSDEFKSRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEKKLPDST	839
Qy	957	EKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIAL	1016

```

      |||||:|||||:|||||
Db      840 EKEDRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIAL 899

Qy      1017 ALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE 1076
      |||||
Db      900 ALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKE 959

Qy      1077 LRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHY 1136
      |||||
Db      960 LRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHY 1019

Qy      1137 LGLANKSVKDAMAKIQAKIPGLKRKAD 1163
      |||||
Db      1020 LGLANKSVKDAMAKIQAKIPGLKRKAE 1046

```

#### RESULT 4

Q8IUA4

```

ID   Q8IUA4          PRELIMINARY;          PRT;   986 AA.
AC   Q8IUA4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RNT4 (RTN4).
GN   RTN4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=22376540; PubMed=12488097;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   J. Mol. Biol. 325:299-323(2003).
DR   EMBL; AY102285; AAM64244.1; -.
DR   EMBL; AY123245; AAM64249.1; -.
DR   EMBL; AY123246; AAM64250.1; -.
DR   EMBL; AY123247; AAM64251.1; -.
DR   EMBL; AY123248; AAM64252.1; -.
DR   EMBL; AY123249; AAM64253.1; -.

```

DR EMBL; AY123250; AAM64254.1; -.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 SQ SEQUENCE 986 AA; 108449 MW; 0CDE8F647036415A CRC64;

Query Match 62.0%; Score 3627.5; DB 4; Length 986;  
 Best Local Similarity 75.4%; Pred. No. 3.9e-174;  
 Matches 745; Conservative 96; Mismatches 126; Indels 21; Gaps 13;

Qy	195	MDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSVFKEHGYLGNL SAVSSSEGTIE	254
		: :                                    :   :	
Db	1	MDLKEQPGNTISAGQEDFPSVLLETAASLPSLSPLSAASFKEHEYLG NLSTVLPTEGTLQ	60
Qy	255	ETLNEASKELPERATNPFVNRDLAEFSELEYSEMGS SFKGS PKGESAILVENTKEEVIVR	314
		:     :  :                      :    :   :	
Db	61	ENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEMGS SFSVSPKAESAVIVANPREEIIVK	120
Qy	315	SKDKED-LVCSAALHSPQESPVG-----KEDRVVSPEKTMDIFNEMQMSV VAPVREEYAD	368
		:   :     :   :                       :      :	
Db	121	NKDEEEKLVSNILHNQQELPTALT KLVKED EVVSSEKAKDSFNEKRVAVEAPMREEYAD	180
Qy	369	FKPFEQAWEVKDTYEGSRDVLAA----RANVESKVDRKCLED SLEQKSLGKDSEGRNEDA	424
		:     :      :    : :    :       :       :	
Db	181	FKPFERVWEVKDSKEDS-DMLAAGGKIESNLESKV DKKCFADSLEQTNHEKDS ESSNDDT	239
Qy	425	SFPSTPEPVKDSSRAYITCASFT-SATESTTANTFPLEDHTSENKTDEKKIEERKAQII	483
		:             :                        :    :	
Db	240	SFPSTPEGIKDRSGAYITCAPFNPAATESIATNIFPLLGDPTSENKTDEKKIEEKKAQIV	299
Qy	484	TEK-TSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNE	542
		:     :	
Db	300	TEKNTSTKTSNPFLVAAQDSETDYVTTDNLT KVTEEVVANMPEGLTPDLVQEACESELNE	359
Qy	543	ATGTKIAYETKVDLVQTSEAIQESLYPTAQLCP SFEEAEATPSPVLPDIVMEAPLNSLLP	602
		:      :           :           :	
Db	360	VTGTKIAYETKMDLVQTSEVMQESLYPAAQLCP SFEESEATPSPVLPDIVMEAPLNSAVP	419
Qy	603	SAGASVVQPSVSPLEAPPPVSYSIKLEPENPPPYEEAMNVALKAL-GTKEGIKEPESFN	661
		:          : :         : :  :          :	
Db	420	SAGASVIQPSSSPLEA-SSVNYESIKHEPENPPPYEEAMSVSLKKVSGIKEEIKEPENIN	478
Qy	662	AAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESEP	721
		:              :    :    :   :   :       :	
Db	479	AALQETEAPYISIACDLIKETKLSAEPAPDFSDYSEMAKVEQVPVDHSELVEDSSPDSEP	538
Qy	722	VDLFSDDSIPEVPQTQEEAVMLMKESLTEVS-ETVAQHK-EERLSASPQELGKPYLESFQ	779
		:     :     :         : :   : :             :	
Db	539	VDLFSDDSI PDVPQKQDETVM LVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFK	598
Qy	780	PNLHSTKDA-ASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFS DSSP	838
		:  :    :     :           : :          : :	
Db	599	LSLDNTKDTLLPDEVSTLSKKEKIP LQMEELSTAVYSNDDL FISK EAQIRETETFS DSSP	658
Qy	839	IEIIDFPTFVSAKDDS-PKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNI	897
		: :        :	

Db 659 IEIIDEFPTLISSKTDSEFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKN I 718

Qy 898 YPK--DEVHVSDEFSENSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPD 955  
 || ::| |::||| | : || : | :||| | | : ||| | | |||||

Db 719 QPKVEEKISFSDDFSNGSATSKVLLLPDVSALATQAEIESIVKPKVLVKEAEKKLPD 778

Qy 956 TEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIA 1015  
 ||||| | : |||||

Db 779 TEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIA 838

Qy 1016 LALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIK 1075  
 |||||

Db 839 LALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIK 898

Qy 1076 ELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDH 1135  
 ||||| : ||||| |||

Db 899 ELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDH 958

Qy 1136 YLGLANKSVKDAMAKIQAKIPGLKRRKAD 1163  
 ||||| : |||||

Db 959 YLGLANKNVKDAMAKIQAKIPGLKRRKAE 986

# RESULT 5

## Q7TNB7

ID Q7TNB7 PRELIMINARY; PRT; 720 AA.

AC Q7TNB7;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC056373; AAH56373.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 720 AA; 77435 MW; 80AB78728F16EAB2 CRC64;

Query Match 56.4%; Score 3299.5; DB 11; Length 720;  
 Best Local Similarity 90.3%; Pred. No. 7.9e-158;  
 Matches 654; Conservative 22; Mismatches 39; Indels 9; Gaps 4;

Qy	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK	60
Db	1	MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEELEVLERK	59
Qy	61	PAAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP	120
Db	60	PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETFLFALP	180
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAAPPSTPAAPKRRGSGSVDETFLFALP	176
Qy	181	AASEPVIPSSAEKIMDLMEQPNGTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	177	AASEPVIPSSAEKIMDLKEQPNGTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	236
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRLAEFSELEYSEMGSSEFKGSPKGES	300
Db	237	GNLSAVASTEGTIEETLNEASRELPERATNPFVNRESAEFSVLEYSEMGSSEFKGSPKGES	296
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESP-----VGKEDRVVSPKTMDFNEMQ	355
		:	
Db	297	AMLVENTKEEVIVRSKDKEDLVCSAALHNPQESPATLTKVVKEDGVMSPEKTMDFNEMK	356
Qy	356	MSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDKCLEDSLEQKSLGK	415
Db	357	MSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANMESKVDKKCFEDSLEQKGHGK	416
Qy	416	DSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKI	475
Db	417	DSESRNENASFPRTPELVKDGSRAYITCDSFSSATESTAANIFPVLEDHTSENKTDEKKI	476
Qy	476	EERKAQIITEKTSNPKTSNPFVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEA	535
Db	477	EERKAQIITEKTSNPKTSNPFVVAIHQDSEADYVTTDNLSKVTEAVVATMPEGLTPDLVQEA	536
Qy	536	CESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEA	595
Db	537	CESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEA	596
Qy	596	PLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIK	655

```

Db          597 PLNSLLPSTGASVAQPSASPLEVSPVSYDGIKLEPENPPPYEEAMSVALKTSDSKEEIK 656
Qy          656 EPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDS 715
          ||||| || |||||||||||||||||||||||||:|||||||||||||||:| |||:|
Db          657 EPESFNAAAQEAAPYISIACDLIKETKLSTEPSPEFSNYSEIAKFEKSVDPHCELVDDS 716
Qy          716 SPES 719
          ||||
Db          717 SPES 720

```

08K290

```

ID      Q8K290                PRELIMINARY;          PRT;    639 AA.
AC      Q8K290;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein.
GN      RTN4.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RL      Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; BC032192; AAH32192.1; -.
DR      MGD; MGI:1915835; Rtn4.
DR      GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR      GO; GO:0007399; P:neurogenesis; IDA.
DR      InterPro; IPR003388; Reticulon.
DR      Pfam; PF02453; Reticulon; 1.
DR      PROSITE; PS50845; RETICULON; 1.
KW      Hypothetical protein.
SQ      SEQUENCE    639 AA;   70312 MW;   309A19DA37603F11 CRC64;

```

Query Match 50.0%; Score 2926; DB 11; Length 639;  
Best Local Similarity 91.6%; Pred. No. 4e-139;  
Matches 588; Conservative 20; Mismatches 30; Indels 4; Gaps 2;

Qy	523	MPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEA	582
Db	1	MPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESIYPTAQLCPSFEEAEA	60
Qy	583	TPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYSYDSIKLEPENPPPYEEAMN	642
Db	61	TPSPVLPDIVMEAPLNSLLPSTGASVAQPSASPLEVSPVSYDGIKLEPENPPPYEEAMS	120
Qy	643	VALKALGTKEGIKEPESFNAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFE	702
Db	121	VALKTSDAKEEIKEPESFNAAAQEAEEAPYISIIACDLIKETKLSTEPSPEFSNYSEIAKFE	180
Qy	703	KSVPEHAELVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHK-EE	761
		:	
Db	181	KSVDPHCELVDSSPESEPVDFLSDDSIPEVPQOTQEEAVMLMKESLTEVSETVTQHKHKE	240

Query Match 44.6%; Score 2610; DB 11; Length 578;







Accession	Start	End	Sequence	Length
Db	188	188	-----	187
Qy	361	420	PVREEYADFKPFQAWVEKDTYEGSRDVLAAARANVESKVDRKCLEDSEGR	420
Db	188	188	-----	187
Qy	421	480	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	188	188	-----	187
Qy	481	540	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	188	188	-----	187
Qy	541	600	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPINSL	600
Db	188	188	-----	187
Qy	601	660	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEAMNVALKALGTKEGIKEPESF	660
Db	188	188	-----	187
Qy	661	720	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	188	188	-----	187
Qy	721	780	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	188	188	-----	187
Qy	781	840	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	188	188	-----	187
Qy	841	900	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	188	188	-----	187
Qy	901	960	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960
Db	188	188	-----	187
Qy	961	1020	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	188	232	-----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	232
Qy	1021	1080	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	233	292	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	292
Qy	1081	1140	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Db	293	352	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	352
Qy	1141	1163	NKSVKDAMAKIQAKIPGLKRRAD	1163

353 NKSVDAMAKIQAKIPGLKRKAE 375

Q8BH78

Query Match 22.5%; Score 1314.5; DB 11; Length 356;  
Best Local Similarity 29.9%; Pred. No. 2e-58;  
Matches 348; Conservative 2; Mismatches 6; Indels 807; Gaps 4;

QY 1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEEEEDDEDLEEEVLERK 60  
||| ||| : |||  
Db 1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEE-EEEDDEDLEEEVLERK 59

---

Qy 61 PAAGLSAAAVPPAAAAPLLDFFSDSVPPAPRGPLPAAPPAAPEROPSWERSPAAPAPSLP 120

Db	60		118
		PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPPTAPERQPSWERSPAASAPSLP	
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPSTPAAPKRRGSGSVDETLFALP	180
Db	119		
		PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPPSTPAAPKRRGSGSV-----	168
Qy	181	AASEFVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	169	-----	168
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Db	169	-----	168
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Db	169	-----	168
Qy	361	PVREEYADFKPFQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSELEQKSLGKDSEGR	420
Db	169	-----	168
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Db	169	-----	168
Qy	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	169	-----	168
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	169	-----	168
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	169	-----	168
Qy	661	NAAVQETEAPYISIIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	169	-----	168
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	169	-----	168
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	169	-----	168
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	169	-----	168
Qy	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKED	960

```

Db          169 ----- 168
Qy          961 RLSAVLSAELSKTSVVDLLYWDRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS 1020
              |||||||||||||||||||||||||||||||||||||||
Db          169 -----VVDLLYWDRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS 213
Qy          1021 VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL 1080
              |||||||||||||||||||||||||||||||||||||||
Db          214 VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL 273
Qy          1081 FLVDDLVDLSLKFAVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA 1140
              |||||||||||||||||||||||||||||||||||||||
Db          274 FLVDDLVDLSLKFAVLMWVFTYVVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLA 333
Qy          1141 NKSVDAMAKIQAKIPGLKRRAD 1163
              |||||||||||||||||||:
Db          334 NKSVDAMAKIQAKIPGLKRRAE 356

```

RESULT 10

Q8K3G7

```

ID   Q8K3G7          PRELIMINARY;      PRT;   357 AA.
AC   Q8K3G7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nogo-B.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Jin W., Li R., Long M., Shen J., Ju G.;
RT   "Cloning and expression of the mouse Nogo-B protein.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY114153; AAM77069.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   357 AA;  38566 MW;  73BB3D17DFDBDF15 CRC64;

```

```

Query Match          22.3%;  Score 1304;  DB 11;  Length 357;
Best Local Similarity 29.9%;  Pred. No. 6.6e-58;
Matches 348;  Conservative 2;  Mismatches 6;  Indels 808;  Gaps 5;

```

```

Qy          1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60
              ||||||||||| |||||||||||||||||||||:| |||||||||||
Db          1 MEDIDQSSLVSSSADSPRPPPAFKYQFVTEPEDEEDEEDEE-EEDEDEDLEEELEVLERK 59
Qy          61 PAAGLSAAAVPPAAAAPLLDFFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120
              ||||| || ||||||||||||||||||||||| ||||||||||| ||||

```

Db	60	PAAGLSAAPVPP-AAAPLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAASAPSLP	118
Qy	121	PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPSTPAAPKRRGSGSVDETLFALP	180
Db	119		
Db	119	PAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAPPSTPAAPKRRGSGSV-----	168
Qy	181	AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL	240
Db	169	-----	168
Qy	241	GNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGSPKGES	300
Db	169	-----	168
Qy	301	AILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQMSVVA	360
Db	169	-----	168
Qy	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEQKSLGKDSEGR	420
Db	169	-----	168
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKTDEKKIEERKA	480
Db	169	-----	168
Qy	481	QIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	169	-----	168
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	169	-----	168
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	169	-----	168
Qy	661	NAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	169	-----	168
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	169	-----	168
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	169	-----	168
Qy	841	IIDEFPFTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	169	-----	168
Qy	901	DEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	169	-----	168

QY 961 RSLSAVLSAELSKTSVVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIALALL 1019  
 Db 169 -----VVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIALALL 213

QY 1020 SVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRR 1079  
 Db 214 SVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRR 273

QY 1080 LFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGL 1139  
 Db 274 LFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGL 333

QY 1140 ANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 334 ANKSVKDAMAKIQAKIPGLKRKAE 357

RESULT 11

Q96B16

ID Q96B16 PRELIMINARY; PRT; 392 AA.  
 AC Q96B16;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein (RTN4).  
 GN RTN4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Oertle T., van der Putten H., Schwab M.E.;  
 RT "Genomic Structure and Functional Characterization of the Promoter  
 Structures of Human and Mouse Nogo/Rtn-4.";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Oertle T., Schwab M.E.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Van der Putten H.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22376540; PubMed=12488097;  
 RA Oertle T., van der Putten H., Schwab M.E.;  
 RT "Genomic Structure and Functional Characterization of the Promoter  
 Structures of Human and Mouse Nogo/Rtn-4.";  
 RL J. Mol. Biol. 325:299-323(2003).



DR EMBL; BC016165; AAH16165.1; -.  
 DR EMBL; AY102285; AAM64242.1; -.  
 DR EMBL; AY102278; AAM64247.1; -.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 392 AA; 42274 MW; D7B2AA5E839E58AD CRC64;

Query Match 21.9%; Score 1283.5; DB 4; Length 392;  
 Best Local Similarity 29.4%; Pred. No. 8.1e-57;  
 Matches 346; Conservative 11; Mismatches 20; Indels 801; Gaps 7;

Qy	1	MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEEVLERK	60
		:                                   :	
Db	1	MEDLDQSPVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEEEVLERK	58
Qy	61	PAAGLSAAAVP---PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P	115
		:   :                   :    :	
Db	59	PAAGLSAAPVPTAPAAGAPLMDFGNDVFPAPRGPLPAAPPVAPERQPSWDPSVSSSTVP	118
Qy	116	APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR	166
Db	119	APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEFVWTPPAPAPAAPPSTPAAPKR	178
Qy	167	RG-SGSVDETLFALPAASEPVIPISSAEKIMDLMEQPGNTVSSSGQEDFPSVLLETAASLPS	225
Db	179	RGSSGSVDETLFALPAASEPVIRSSA-----	204
Qy	226	LSPLSTVSFKEHGYLGNL SAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEY	285
Db	205	-----	204
Qy	286	SEMGSFSGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPE	345
Db	205	-----	204
Qy	346	KTMDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLE	405
Db	205	-----	204
Qy	406	DSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHT	465
Db	205	-----	204
Qy	466	SENKTDEKKIEERKAQIIITEKTSNPNFLVAVQDSEADYVTTDTLSKVTEAAVSNMPE	525
Db	205	-----	204
Qy	526	GLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPS	585
Db	205	-----	204
Qy	586	PVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVAL	645
Db	205	-----	204

Qy 646 KALGTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKSV 705  
 Db 205 ----- 204  
 Qy 706 PEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSA 765  
 Db 205 ----- 204  
 Qy 766 SPQELGKPYLESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFNNTAIYSNDDLSSKED 825  
 Db 205 ----- 204  
 Qy 826 KIKESETFSDSSPIELIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPC 885  
 Db 205 ----- 204  
 Qy 886 LELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLT 945  
 Db 205 ----- 204  
 Qy 946 KEAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF 1005  
 Db 205 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVF 234  
 Qy 1006 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 1065  
 Db 235 SIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNS 294  
 Qy 1066 ALGHVNSTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVI 1125  
 Db 295 ALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI 354  
 Qy 1126 YERHQVQIDHYLGLANKSVKDAMAKIQAKIPGLKRKAD 1163  
 Db 355 YERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 392

RESULT 12

Q7YRW9

ID Q7YRW9 PRELIMINARY; PRT; 184 AA.  
 AC Q7YRW9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE RTN4w (Fragment).  
 GN RTN4.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22715887; PubMed=12832288;  
 RA Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;  
 RT "A reticular rhapsody: phylogenic evolution and nomenclature of the



```

Qy      985 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKQVIQAIQKSDEGHPF 1044
      :|||||||||||||||||||||||||||||||||||||:|||||||
Db      1 MKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYRGVIQAIQKSDEGHPF 60

Qy     1045 RAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFVLMWVFTYVGA 1104
      ||||||||||||||||||||||||| |||||||||:|||||||
Db     61 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDLSLRFVLMWVFTYVGA 120

Qy     1105 LFNGLTLLILALISLFSIPVIYERHOVIDHYLGLANKSVKDAMAKIOAKIPGLKRKAD 1163

```

|||||:|||||:|||||:|||||:
Db 121 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 179

RESULT 15

Q8K4S4

ID Q8K4S4 PRELIMINARY; PRT; 780 AA.
AC Q8K4S4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Reticulon 1A.
GN RTN1 OR RTN-1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI\_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RA Hirata T., Nomura T., Takagi Y., Sato Y., Tomioka N., Fujisawa H.,
RA Osumi N.;
RT "Mosaic development of the olfactory cortex with Pax6-dependent and -
RT independent components.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB074899; BAB96551.1; -.
DR MGD; MGI:1933947; Rtn1.
DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR InterPro; IPR001951; Histone\_H4.
DR InterPro; IPR003388; Reticulon.
DR Pfam; PF02453; Reticulon; 1.
DR PROSITE; PS00047; HISTONE\_H4; 1.
DR PROSITE; PS50845; RETICULON; 1.
SQ SEQUENCE 780 AA; 83504 MW; 545F5638C576A069 CRC64;

Query Match 13.5%; Score 792; DB 11; Length 780;
Best Local Similarity 32.5%; Pred. No. 1.1e-31;
Matches 242; Conservative 105; Mismatches 260; Indels 138; Gaps 27;

Qy 517 EAAVSNMPEGLTPDLVQEACESELNEATG-----TKIAYETKVDLVQTSEAIQ-ESLY 568
| | : | | : : | | : | : : | : | : : | | : :
Db 76 ETASTGM--AAVPDALDHSPSSTLKDGEGACYTSLISDVCYPPREDSAYFTGILQKENGH 133
Qy 569 PTAQLCPSPFEEAEATPSPVLPDIVMEAP---LNS-----LLPSAGASVVQPSVSPLEAPP 620
| | | | | | | | : | | : | : | : | : | : | :
Db 134 ITTSESP--EEPE-TPGPSLPEVPGMEPQGLLSSDSGIEMTPAESTEVNKILADPLDQMK 190
Qy 621 PVSYSIDIKLE-----PENPPPYEEAMNVALKALGTK-----EGIKEPESFNAAV---- 664
: | | : | | | : : | | : | | : | | : | |
Db 191 AEAYKYIDITRPQEAKGQEEQHPGLEDKDLDFKDKGTEVSTKAEGVRAPNQ-PAPVEGKL 249
Qy 665 -----QETEAPYISIIACDLIKETKLSTEPSPDFSNYSEI-----AKFEKSVPEHAELV 712
: | | | | | : : : | : | : | : : | : |
Db 250 IKDHLFEESTFAPYIDELSD--EQHRVSLVTAPVKITLTEIEPPLMTATQETIPEKQDLC 307
Qy 713 EDSSPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHK-----EERLSASP 767
| | : | | : | : : | | | : | | : | : |



OM protein - protein search, using sw model

Run on: September 29, 2004, 18:04:48 ; Search time 16.6724 Seconds  
(without alignments)  
3632.211 Million cell updates/sec

Title: US-09-830-972-2  
Perfect score: 5848  
Sequence: 1 MEDIDQSSSLVSSSTDSPRP.....VKDAMAKIQAKIPGLKRKAD 1163

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	5848	100.0	1163	1	RTN4_RAT	Q9jkl1 rattus norv
2	4403.5	75.3	1192	1	RTN4_HUMAN	Q9nqc3 homo sapien
3	917	15.7	199	1	RTN4_MOUSE	Q99p72 mus musculu
4	801	13.7	777	1	RTN1_RAT	Q64548 rattus norv
5	789.5	13.5	776	1	RTN1_HUMAN	Q16799 homo sapien
6	625.5	10.7	236	1	RTN3_HUMAN	O95197 homo sapien
7	625.5	10.7	237	1	RTN3_MOUSE	Q9es97 mus musculu
8	509	8.7	471	1	RTN2_MOUSE	O70622 mus musculu
9	490	8.4	545	1	RTN2_HUMAN	O75298 homo sapien
10	308.5	5.3	5147	1	PCLO_HUMAN	Q9y6v0 homo sapien
11	300	5.1	5120	1	PCLO_CHICK	Q9pu36 gallus gall
12	295.5	5.1	865	1	CPN_DROME	Q02910 drosophila
13	291.5	5.0	2459	1	MAPB_RAT	P15205 rattus norv
14	288	4.9	2468	1	MAPB_HUMAN	P46821 homo sapien
15	284	4.9	2464	1	MAPB_MOUSE	P14873 mus musculu
16	282.5	4.8	2805	1	MAPA_HUMAN	P78559 homo sapien
17	279	4.8	5038	1	PCLO_MOUSE	Q9qyx7 mus musculu

18	277	4.7	3644	1	MINT_MOUSE	Q62504	mus musculu
19	275.5	4.7	3924	1	ANK2_HUMAN	Q01484	homo sapien
20	275	4.7	1972	1	P531_HUMAN	Q12888	homo sapien
21	272.5	4.7	3664	1	MINT_HUMAN	Q96t58	homo sapien
22	272	4.7	5085	1	PCLO_RAT	Q9jks6	rattus norv
23	267	4.6	1781	1	AK12_HUMAN	Q02952	homo sapien
24	265.5	4.5	1828	1	MAP2_MOUSE	P20357	mus musculu
25	261.5	4.5	1087	1	NFH_MOUSE	P19246	mus musculu
26	259.5	4.4	1026	1	NFH_HUMAN	P12036	homo sapien
27	259.5	4.4	2404	1	SON_MOUSE	Q9qx47	mus musculu
28	258.5	4.4	1861	1	MAP2_RAT	P15146	rattus norv
29	257.5	4.4	2715	1	MLL4_HUMAN	Q9umn6	homo sapien
30	256	4.4	1723	1	AIM1_HUMAN	Q9y4k1	homo sapien
31	255.5	4.4	1140	1	YM96_YEAST	Q04893	saccharomyc
32	254.5	4.4	6632	1	UN89_CAEEL	O01761	caenorhabdi
33	254	4.3	1616	1	P200_MYCGE	Q49429	mycoplasma
34	254	4.3	3381	1	PGCV_BOVIN	P81282	bos taurus
35	253	4.3	3421	1	TEGU_HSVB	P28955	equine herp
36	252.5	4.3	1411	1	TCOF_HUMAN	Q13428	homo sapien
37	252	4.3	2738	1	PGCV_RAT	Q9erb4	rattus norv
38	250.5	4.3	4377	1	ANK3_HUMAN	Q12955	homo sapien
39	248.5	4.2	1189	1	YJH6_YEAST	P47035	saccharomyc
40	248	4.2	1827	1	MAP2_HUMAN	P11137	homo sapien
41	245.5	4.2	1001	1	IF2_SYNY3	P72689	synechocyst
42	244	4.2	1338	1	ACIN_MOUSE	Q9jix8	mus musculu
43	241.5	4.1	8545	1	ANC1_CAEEL	Q9n4m4	caenorhabdi
44	238.5	4.1	831	1	NFH_RAT	P16884	rattus norv
45	238.5	4.1	1341	1	ACIN_HUMAN	Q9ukv3	homo sapien

#### ALIGNMENTS

##### RESULT 1

##### RTN4\_RAT

ID RTN4\_RAT STANDARD; PRT; 1163 AA.  
AC Q9JK11; Q9JK10; Q9R0D9; Q9WUE9; Q9WUF0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)  
DE (Glut4 vesicle 20 kDa protein).  
GN RTN4 OR NOGO.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.  
RC STRAIN=Sprague-Dawley; TISSUE=Adipocyte;  
RX MEDLINE=99249816; PubMed=10231557;  
RA Morris N.J., Ross S.A., Neveu J.M., Lane W.S., Lienhard G.E.;  
RT "Cloning and characterization of a 22 kDa protein from rat adipocytes:  
RT a new member of the reticulon family."  
RL Biochim. Biophys. Acta 1450:68-76(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).



RX MEDLINE=20129258; PubMed=10667796;  
 RA Chen M.S., Huber A.B., Van der Haar M.E., Frank M., Schnell L.,  
 RA Spillmann A.A., Christ F., Schwab M.E.;  
 RT "Nogo-A is a myelin-associated neurite outgrowth inhibitor and an  
 RT antigen for monoclonal antibody IN-1.";  
 RL Nature 403:434-439(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).  
 RC STRAIN=Wistar Kyoto; TISSUE=Vascular smooth muscle;  
 RA Ito T., Schwartz S.M.;  
 RT "Cloning of a member of the reticulon gene family in rat: one of two  
 RT minor splice variants.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE=22033691; PubMed=12037567;  
 RA GrandPre T., Li S., Strittmatter S.M.;  
 RT "Nogo-66 receptor antagonist peptide promotes axonal regeneration.";  
 RL Nature 417:547-551(2002).  
 CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help  
 CC block the regeneration of the nervous central system in adults (By  
 CC similarity).  
 CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-x1 and Bcl-2 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the  
 CC membrane of the endoplasmic reticulum through 2 putative  
 CC transmembrane domains (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1; Synonyms=Nogo-A, NI-220-250;  
 CC IsoId=Q9JK11-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Nogo-B, Foocen-M1;  
 CC IsoId=Q9JK11-2; Sequence=VSP\_005658;  
 CC Name=3; Synonyms=Nogo-C, VP20;  
 CC IsoId=Q9JK11-3; Sequence=VSP\_005656, VSP\_005657;  
 CC Name=4; Synonyms=Foocen-M2;  
 CC IsoId=Q9JK11-4; Sequence=VSP\_005659;  
 CC -!- TISSUE SPECIFICITY: Isoforms 1, 2 and 3 are present in optic  
 CC nerve, spinal cord and cerebral cortex. Isoforms 1 and 2 are  
 CC present in dorsal root ganglion, sciatic nerve and PC12 cells  
 CC after longer exposure. Isoforms 2 and 3 are detected in kidney,  
 CC cartilage, skin, lung and spleen. Isoform 3 is expressed at high  
 CC level in skeletal muscle. In adult animals isoform 1 is expressed  
 CC mainly in the nervous system.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF051335; AAF01564.1; -.  
 DR EMBL; AJ242961; CAB71027.1; -.  
 DR EMBL; AJ242962; CAB71028.1; -.

DR EMBL; AJ242963; CAB71029.1; -.  
 DR EMBL; AF132045; AAD31019.1; -.  
 DR EMBL; AF132046; AAD31020.1; -.  
 DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.  
 DR GO; GO:0005635; C:nuclear membrane; ISS.  
 DR GO; GO:0005515; F:protein binding; ISS.  
 DR GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.  
 DR GO; GO:0030517; P:negative regulation of axon extension; ISS.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane.  
 FT DOMAIN 1 989 CYTOPLASMIC (Potential).  
 FT TRANSMEM 990 1010 POTENTIAL.  
 FT DOMAIN 1011 1104 LUMENAL (Potential).  
 FT TRANSMEM 1105 1125 POTENTIAL.  
 FT DOMAIN 1126 1163 CYTOPLASMIC (Potential).  
 FT DOMAIN 976 1163 RETICULON.  
 FT DOMAIN 33 46 POLY-GLU.  
 FT DOMAIN 73 76 POLY-ALA.  
 FT DOMAIN 140 145 POLY-PRO.  
 FT VARSPLIC 1 964 Missing (in isoform 3).  
 FT /FTId=VSP\_005656.  
 FT VARSPLIC 965 975 AVLSAELSKTS -> MDGQKKHWKDK (in isoform 3).  
 FT /FTId=VSP\_005657.  
 FT VARSPLIC 173 975 Missing (in isoform 2).  
 FT /FTId=VSP\_005658.  
 FT VARSPLIC 192 975 Missing (in isoform 4).  
 FT /FTId=VSP\_005659.  
 FT CONFLICT 1130 1131 MISSING (IN REF. 3; AAD31020).  
 SQ SEQUENCE 1163 AA; 126386 MW; 8CB894B09E94F0B6 CRC64;

Query Match 100.0%; Score 5848; DB 1; Length 1163;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-213;  
 Matches 1163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEDIDQSSLVSSSTDSPRPPPAFKYQFVTEPEDEEDEEEEEDEEDEDLEEELEVLERK 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MEDIDQSSLVSSSTDSPRPPPAFKYQFVTEPEDEEDEEDEEEEEDEEDEDLEEELEVLERK 60  
  
 Qy 61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 PAAGLSAAAVPPAAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAAPAPSLP 120  
  
 Qy 121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALP 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 PAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPSTPAAPKRRGSGSVDETLFALP 180  
  
 Qy 181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 AASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYL 240  
  
 Qy 241 GNLSAVSSSEGTIEETLNEASKELPERATNPVNRDLAEFSELEYSEMGSSFKGSPKGES 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 GNLSAVSSSEGTIEETLNEASKELPERATNPVNRDLAEFSELEYSEMGSSFKGSPKGES 300

Qy	301	AILVENTKEEVIVRSKDKEDIVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360
Db	301	AILVENTKEEVIVRSKDKEDIVCSAALHSPQESPVGKEDRVVSPEKTMDFNEMQMSVVA	360
Qy	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEQKSLGKDSEGR	420
Db	361	PVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVESKVDRKCLEDSEQKSLGKDSEGR	420
Qy	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Db	421	NEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKIEERKA	480
Qy	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Db	481	QIITEKTS PKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESEL	540
Qy	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Db	541	NEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSL	600
Qy	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Db	601	LPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESF	660
Qy	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Db	661	NAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESE	720
Qy	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Db	721	PVDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYLESFQP	780
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Qy	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Db	841	IIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPK	900
Qy	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Db	901	DEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKED	960
Qy	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Db	961	RSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS	1020
Qy	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Db	1021	VTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL	1080
Qy	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Db	1081	FLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLA	1140
Qy	1141	NKSVKDAMAKIQAKIPGLKRRAD	1163

Db

|||||  
1141 NKS VKDAMAKIQA KIPGLKRKAD 1163

RESULT 2

RTN4\_HUMAN

ID RTN4\_HUMAN STANDARD; PRT; 1192 AA.  
AC Q9NQC3; Q94962; Q9BXG5; Q9H212; Q9H3I3; Q9UQ42; Q9Y293; Q9Y2Y7;  
AC Q9Y5U6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)  
DE (Neuroendocrine-specific protein) (NSP) (Neuroendocrine specific  
DE protein C homolog) (RTN-x) (Reticulon 5) (My043 protein).  
GN RTN4 OR NOGO OR ASY OR KIAA0886.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=20129242; PubMed=10667780;  
RA Prinjha R., Moore S.E., Vinson M., Blake S., Morrow R., Christie G.,  
RA Michalovich D., Simmons D.L., Walsh F.S.;  
RT "Inhibitor of neurite outgrowth in humans.";  
RL Nature 403:383-384(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Brain;  
RX MEDLINE=21010696; PubMed=11126360;  
RA Tagami S., Eguchi Y., Kinoshita M., Takeda M., Tsujimoto Y.;  
RT "A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on  
RT endoplasmic reticulum and reduces their anti-apoptotic activity.";  
RL Oncogene 19:5736-5746(2000).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=20237542; PubMed=10773680;  
RA Yang J., Yu L., Bi A.D., Zhao S.-Y.;  
RT "Assignment of the human reticulon 4 gene (RTN4) to chromosome  
RT 2p14-->2p13 by radiation hybrid mapping.";  
RL Cytogenet. Cell Genet. 88:101-102(2000).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 4).  
RA Jin W.-L., Ju G.;  
RT "Developmentally-regulated alternative splicing in a novel Nogo-A.";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).  
RC TISSUE=Placenta, and Skeletal muscle;  
RA Ito T., Schwartz S.M.;  
RT "Cloning of a member of the reticulon gene family in human.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Fibroblast;  
RA Yutsudo M.;

RT "Isolation of a cell death-inducing gene."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Pituitary;  
 RA Song H., Peng Y., Zhou J., Huang Q., Dai M., Mao Y.M., Yu Y., Xu X.,  
 RA Luo B., Hu R., Chen J.;  
 RT "Human neuroendocrine-specific protein C (NSP) homolog gene."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,  
 RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,  
 RA Yu J., Han L.H.;  
 RT "Novel human cDNA clone with function of inhibiting cancer cell  
 RT growth."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=99156230; PubMed=10048485;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 5:355-364(1998).  
 RN [10]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).  
 RC TISSUE=Brain, Ovary, Pancreas, Placenta, and Skeletal muscle;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [11]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,  
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,  
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;

RT "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560(2000).  
 RN [12]  
 RP SEQUENCE OF 482-1192 FROM N.A. (ISOFORM 1/4).  
 RC TISSUE=Brain;  
 RA Mao Y.M., Xie Y., Zheng Z.H.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [13]  
 RP SEQUENCE OF 186-1192 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Testis;  
 RA Sha J.H., Zhou Z.M., Li J.M.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [14]  
 RP TOPOLOGY.  
 RC TISSUE=Brain;  
 RX MEDLINE=20129259; PubMed=10667797;  
 RA GrandPre T., Nakamura F., Vartanian T., Strittmatter S.M.;  
 RT "Identification of the Nogo inhibitor of axon regeneration as a  
 RT Reticulon protein.";  
 RL Nature 403:439-444(2000).  
 RN [15]  
 RP FUNCTION.  
 RC TISSUE=Brain;  
 RX MEDLINE=21069055; PubMed=11201742;  
 RA Fournier A.E., Grandpre T., Strittmatter S.M.;  
 RT "Identification of a receptor mediating Nogo-66 inhibition of axonal  
 RT regeneration.";  
 RL Nature 409:341-346(2001).  
 RN [16]  
 RP REVIEW.  
 RX MEDLINE=21888956; PubMed=11891768;  
 RA Ng C.E.L., Tang B.L.;  
 RT "Nogos and the Nogo-66 receptor: factors inhibiting CNS neuron  
 RT regeneration.";  
 RL J. Neurosci. Res. 67:559-565(2002).  
 CC -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help  
 CC block the regeneration of the nervous central system in adults.  
 CC Isoform 2 reduces the anti-apoptotic activity of Bcl-xl and Bcl-2.  
 CC This is likely consecutive to their change in subcellular  
 CC location, from the mitochondria to the endoplasmic reticulum,  
 CC after binding and sequestration.  
 CC -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum. Anchored to the membrane of the endoplasmic reticulum  
 CC through 2 putative transmembrane domains.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1; Synonyms=RTN 4A, Nogo-A, RTN-xL;  
 CC IsoId=Q9NQC3-1; Sequence=Displayed;  
 CC Name=2; Synonyms=RTN 4B, Nogo-B, RTN-xS, Foocen-M;  
 CC IsoId=Q9NQC3-2; Sequence=VSP\_005655;  
 CC Name=3; Synonyms=RTN 4C, Nogo-C, Foocen-S;  
 CC IsoId=Q9NQC3-3; Sequence=VSP\_005652, VSP\_005653;  
 CC Name=4;  
 CC IsoId=Q9NQC3-4; Sequence=VSP\_005654;

CC -!- TISSUE SPECIFICITY: Isoform 1 is specifically expressed in brain  
CC and testis and weakly in heart and skeletal muscle. Isoform 2 is  
CC widely expressed excepted for the liver. Isoform 3 is expressed in  
CC brain, skeletal muscle and adipocytes. Isoform 4 is testis-  
CC specific.  
CC -!- SIMILARITY: Contains 1 reticulon domain.  
CC -!- CAUTION: Ref.11 sequence differs from that shown due to  
CC frameshifts in positions 1149 and 1156.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AJ251383; CAB99248.1; -.  
DR EMBL; AJ251384; CAB99249.1; -.  
DR EMBL; AJ251385; CAB99250.1; -.  
DR EMBL; AB040462; BAB18927.1; -.  
DR EMBL; AB040463; BAB18928.1; -.  
DR EMBL; AF148537; AAG12176.1; -.  
DR EMBL; AF148538; AAG12177.1; -.  
DR EMBL; AF087901; AAG12205.1; -.  
DR EMBL; AF320999; AAG40878.1; -.  
DR EMBL; AF132047; AAD31021.1; -.  
DR EMBL; AF132048; AAD31022.1; -.  
DR EMBL; AB015639; BAA83712.1; -.  
DR EMBL; AF077050; AAD27783.1; -.  
DR EMBL; AF177332; AAG17976.1; -.  
DR EMBL; AB020693; BAA74909.1; -.  
DR EMBL; BC001035; AAH01035.1; -.  
DR EMBL; BC007109; AAH07109.1; -.  
DR EMBL; BC014366; AAH14366.1; -.

Query Match 75.3%; Score 4403.5; DB 1; Length 1192;  
Best Local Similarity 75.9%; Pred. No. 9.1e-159;  
Matches 909; Conservative 104; Mismatches 145; Indels 39; Gaps 20;

Qy 1 MEDIDQSSLVSSSTDSPRPAPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60  
|||:|||| ||||| ||||| ||||| ||||| ||||| |||||:||||| |||||  
Db 1 MEDLDQSPLVSSS-DSPRPQPAFKYQFVREPEDEE-EEEEEEEEDEDEDLEELEVLERK 58  
  
Qy 61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115  
||||||| || ||| |||:| | ||||| ||||| |||||: || : |  
Db 59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSVSSSTVP 118  
  
Qy 116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAE-----PAAPPSTPAAPKR 166  
||| |||| ||||| ||||| ||||| || || ||||| |||||  
Db 119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPAPAPAAPSTPAAPKR 178  
  
Qy 167 RG-SGSVDETLFALPAASEVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPS 225  
|| ||||| ||||| |||| ||| |||||:|:||||| |||||  
Db 179 RGSSGSVDETLFALPAASEVIRSSAEN-MDLKEQPGNTISAGQEDFPSVLLETAASLPS 237  
  
Qy 226 LSPLSTVSFKEHGYLGNLSAVSSSEGTTIEETLNEASKELPERATNPFVNRDLAEFSELEY 285

Db	238	LSPLSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEY	297
Qy	286	SEMGSSFSGSPKGESAILVENTKEEVIVRSKDKED-LVCSAALHSPQESPVG-----KED	339
Db	298	SEMGSSFSVSPKAESAVIVANPREETIVKNKDEEEKLVSNILHNQQELPTALTCLKVKED	357
Qy	340	RVVSPEKTMDFNEMQMSVVAPVREEYADFKPFQAWEVKDTYEGSRDVLAA-----RANV	395
Db	358	EVVSSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDS-DMLAAGGKIESNL	416
Qy	396	ESKVDRCLEDSLEQKSLGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFT-SATESTT	454
Db	417	ESKVDKKCFADSLEQTNHEKDSSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIA	476
Qy	455	ANTFPLLEDHTSENKTDEKKIEERKAQIITEK-TSPKTSNPFLVAVQDSEADYVTTDTLS	513
Db	477	TNIFPLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLT	536
Qy	514	KVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQL	573
Db	537	KVTEEVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQL	596
Qy	574	CPSFEEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPEN	633
Db	597	CPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEA-SSVNYESIKHEPEN	655
Qy	634	PPPYEEAMNVALKAL-GTKEGIKEPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDF	692
Db	656	PPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDF	715
Qy	693	SNYSEIAKFEKSVPEHAELVEDSSPESEPVDFLSDDSIPEVPQTQEEAVMLMKESLTEVS	752
Db	716	SDYSEMAKVEQVPDPHSELVEDSSPDSEPVDFLSDDSIPDVPQKQDETVMVLKESLTETS	775
Qy	753	-ETVAQHK-EERLSASPQELGKPYLESFQPNLHSTKDA-ASNDIPTLTKEKISLQMEEF	809
Db	776	FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEEL	835
Qy	810	NTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDFPTFVSAKDDS-PKLAKEYTDLEVS	868
Db	836	STAVYSNDDLFIKSEAQIRETETFSDDSSPIEIIDFPTLISSKTDSEFSLKAREYTDLEVS	895
Qy	869	DKSEIANIQSGADSLPCLELPCDLSEFKNIYPK--DEVHVSDEFSENRSSVSKASISPSNV	926
Db	896	HKSEIANAPDGAGSLPCTELPHDLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPDVP	955
Qy	927	SALEPQTEMGSIKSKSLTKEAEKKLPDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIK	986
Db	956	SALATQAEIESIVKPKVLVKEAEKKLPDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIK	1015
Qy	987	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1046
Db	1016	KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIAIQKSDEGHPFRA	1075
Qy	1047	YLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALF	1106



```

ID      RTN4 MOUSE          STANDARD;          PRT;          199 AA.
AC      Q99P72; Q9CTE3;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein).
GN      RTN4 OR NOGO.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3T3-L1; TISSUE=Adipocyte;
RA      Coulson A.C., Craggs P.D., Morris N.J.;
RT      "Mouse vp20/RTN4C cDNA.";
RL      Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE OF 170-199 FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Embryo;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA      Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA      Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA      Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
CC      -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC      block the regeneration of the nervous central system in adults (By
CC      similarity).
CC      -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
CC      similarity).
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
CC      membrane of the endoplasmic reticulum through 2 putative

```

[illegible]

RESULT 4

RTN1\_RAT

ID RTN1\_RAT STANDARD; PRT; 777 AA.  
 AC Q64548; Q64547;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Reticulon 1 (Neuroendocrine-specific protein) (S-rex).  
 GN RTN1 OR NSP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).

RC STRAIN=Wistar; TISSUE=Brain cortex;

RX MEDLINE=96386034; PubMed=8793864;

RA Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,

RA Georgiev G.P., Buchman V.L.;

RT "Intracellular compartmentalization of two differentially spliced s-  
 RT rex/NSP mRNAs in neurons.";

RL Mol. Cell. Neurosci. 7:289-303(1996).

CC -!- FUNCTION: May be involved in neuroendocrine secretion or in  
 CC membrane trafficking in neuroendocrine cells.

CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane (By  
 CC similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=RTN1-B; Synonyms=S-RexB;

CC IsoId=Q64548-1; Sequence=Displayed;

CC Name=RTN1-S; Synonyms=S-RexS;

CC IsoId=Q64548-2; Sequence=VSP\_005647, VSP\_005648;

CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND  
 CC PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS  
 CC HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.  
 CC EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL  
 CC TYPES.

CC -!- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE  
 CC HINDBRAIN AND IN E11 IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC  
 CC DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN  
 CC THE HINDBRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB  
 CC DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE  
 CC HINDBRAIN.

CC -!- SIMILARITY: Contains 1 reticulon domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U17604; AAC53046.1; -.

DR EMBL; U17603; AAC53045.1; -.

DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane.  
 FT TRANSMEM 604 624 POTENTIAL.  
 FT TRANSMEM 727 747 POTENTIAL.  
 FT DOMAIN 590 777 RETICULON.  
 FT DOMAIN 610 613 POLY-LEU.  
 FT VARSPLIC 1 569 Missing (in isoform RTN1-S).  
 FT /FTid=VSP\_005647.  
 FT VARSPLIC 570 589 IPGPLGSDLVPPPLPFFNKQK -> MQATADSTKMDCVWSNW  
 FT KSQ (in isoform RTN1-S).  
 FT /FTid=VSP\_005648.  
 SQ SEQUENCE 777 AA; 83001 MW; AF7479C50F28D0AC CRC64;

Query Match 13.7%; Score 801; DB 1; Length 777;  
 Best Local Similarity 32.4%; Pred. No. 2e-23;  
 Matches 258; Conservative 109; Mismatches 264; Indels 166; Gaps 32;

Qy 485 EKTSPKTSNPFLVAVQDSE-----ADYVTTDTL----SKVTEAAVSNMPEGL--TPD 530  
 |: :|| : | | || | | :| : : | : | : ||  
 Db 29 EEATPKGARP---AQQDGEPAWGSGAGAGVVSSRGLCSGPARSPVAMETASTGVAAPD 85

Qy 531 LVQEACESELNEATG-----TKIAYETKVDLVQTSEAIQ-ESLYPTAQLCPSFEEAEA 582  
 : : | : | : | | : | : : | : | : |  
 Db 86 ALDHSSSPTLKDGEACYSLSIDICYPPREDSAYFTGILQKENGHITTSESP---EELG 142

Qy 583 TPSPVLPDIVMEAPLNSLLPSAGASVVQPS----VSPLEAPPPVSYDSIKLE----- 630  
 || | ||:: | : || | : | : | : | : | : |  
 Db 143 TPGPSLPEVPGTEP-HGLLSSDSGIEMTPAESTEVNKLADP---LDQMKAACKYIDIT 198

Qy 631 -----PENPPPYEEAMNVALK----ALGTK-EGIKEPE-----SFNAAVQETE 668  
 | | | : : | : | | : | : | : | : |  
 Db 199 RPQEAKGQEEQSPGLEDKDLDFKDKDSEVSTKPEGVHAPNQSPVEGKLIKDNLFEESTF 258

Qy 669 APYISIACDLIKETKLSTEPSPDFSNYSEI-----AKFEKSVPEHAELVEDSSPESEPV 723  
 |||| | : : : | : | : || : : || : | || : : |  
 Db 259 APYIDELSD--EQHRMSLVTAPVKITLTEIGPPVMTATHETIPEKQDLCLKPSPTVPTV 316

Qy 724 LFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHK-----EERLSASPQELGKPYLESF 778  
 | : | : : | | || | : | : | : | : | : |  
 Db 317 TVSE-----PEDDSPGSVTPPSSGTEPSAAESQGKGSVSEDELIAAIKEAKGLSYETTE 370

Qy 779 QPN-----LHSTKDAASNDIPTLTKEKISLQMEEFNATYISNDDLSSKEDKIKESETF 833  
 | | | : : || : | | || | | | |  
 Db 371 SPRPVGQAADRPKVKARSGLPTIPS-----SLDQEASSAESGDSEIELV 414

Qy 834 SDSSPIEIIDEFPT-----FVSAKDDSPKLAK-EYTDLEVSDKSEIANIQSGADSLPCL 886  
 | : | : | : | | | : | : | : : | : | :  
 Db 415 SE-DPMASEDALPSGYVSFGHVSGPPPPSPASPSIQYSILREEREAEEL-----DSELI 466

Qy 887 ELPCDL-SFKNIYPKDEVH-----VSDEFSENRSSVSKASISPSNVSALEPQTEM--- 935  
 | || | || | | | | : : || : : || :  
 Db 467 E-SCDASSASEESPKREQDSPPMKPGVLDAREETSSRATEERAPSHQGPVEPDPILSFT 525

Qy 936 -----GSIV----KSKSLTKEAE----KKLPSDTEKEDRSLSAVLSAEL-- 971  
 | : | || : | | : | : || : | : ||



```

CC  -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC      membrane trafficking in neuroendocrine cells.
CC  -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
CC  -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=3;
CC      Name=RTN1-A; Synonyms=NSP-A;
CC      IsoId=Q16799-1; Sequence=Displayed;
CC      Name=RTN1-B; Synonyms=NSP-B;
CC      IsoId=Q16799-2; Sequence=VSP_005644;
CC      Name=RTN1-C; Synonyms=NSP-C;
CC      IsoId=Q16799-3; Sequence=VSP_005645, VSP_005646;
CC  -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL AND NEUROENDOCRINE TISSUES
CC      AND CELL CULTURES DERIVED THEREFROM. EXPRESSION OF ISOFORM RTN1-C
CC      IS STRONGLY CORRELATED WITH NEURONAL DIFFERENTIATION.
CC  -!- PTM: Isoforms RTN1-A and RTN1-B are phosphorylated.
CC  -!- SIMILARITY: Contains 1 reticulon domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L10333; AAA59950.1; -.
DR  EMBL; L10334; AAA59951.1; -.
DR  EMBL; L10335; AAA59952.1; -.
DR  PIR; A46583; A46583.
DR  PIR; I60904; I60904.
DR  Genew; HGNC:10467; RTN1.
DR  MIM; 600865; -.
DR  GO; GO:0030176; C:integral to endoplasmic reticulum membrane; TAS.
DR  GO; GO:0004871; F:signal transducer activity; NAS.
DR  GO; GO:0030182; P:neuron differentiation; TAS.
DR  GO; GO:0007165; P:signal transduction; NAS.
DR  InterPro; IPR003388; Reticulon.
DR  Pfam; PF02453; Reticulon; 1.
DR  PROSITE; PS50845; RETICULON; 1.
KW  Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW  Phosphorylation.
FT  TRANSMEM      603      623      POTENTIAL.
FT  TRANSMEM      726      746      POTENTIAL.
FT  DOMAIN        589      776      RETICULON.
FT  DOMAIN        609      612      POLY-LEU.
FT  VARSPLIC       1      420      Missing (in isoform RTN1-B).
FT                                     /FTId=VSP_005644.
FT  VARSPLIC       1      568      Missing (in isoform RTN1-C).
FT                                     /FTId=VSP_005645.
FT  VARSPLIC      569      588      GPGPLGPGAPPPLFLNKQK -> MQATADSTKMDCVWSNW
FT                                     KSQ (in isoform RTN1-C).
FT                                     /FTId=VSP_005646.
SQ  SEQUENCE      776 AA;  83617 MW;  CA5B6232353096FE CRC64;

Query Match          13.5%;  Score 789.5;  DB 1;  Length 776;
Best Local Similarity 31.2%;  Pred. No. 5.3e-23;
Matches 243;  Conservative 115;  Mismatches 252;  Indels 169;  Gaps 28;

```



AC 095197;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Reticulon protein 3 (Neuroendocrine-specific protein-like 2) (NSP-like  
 DE protein II) (NSPLII).  
 GN RTN3 OR NSPL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Retina;  
 RX MEDLINE=99265974; PubMed=10331947;  
 RA Moreira E.F., Jaworski C.J., Rodriguez I.R.;  
 RT "Cloning of a novel member of the reticulon gene family (RTN3): gene  
 RT structure and chromosomal localization to 11q13.";  
 RL Genomics 58:73-81(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Huang X., Zhou Y., Du G., Yuan J., Qiang B.;  
 RT "Cloning and expression analysis of a cDNA encoding a novel  
 RT neuroendocrine-specific protein-like protein 1: NSPL1.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, Eye, and Lymph;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum (Potential).  
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN  
 CC BRAIN. THREE TIMES MORE ABUNDANT IN MACULA THAN IN PERIPHERAL  
 CC RETINA.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF059524; AAC99319.1; -.  
DR EMBL; AF059529; AAD20951.1; -.  
DR EMBL; AF059525; AAD20951.1; JOINED.  
DR EMBL; AF059526; AAD20951.1; JOINED.  
DR EMBL; AF059527; AAD20951.1; JOINED.  
DR EMBL; AF059528; AAD20951.1; JOINED.  
DR EMBL; AF119297; AAD26810.1; -.  
DR EMBL; BC000634; AAH00634.1; -.  
DR EMBL; BC010556; AAH10556.1; -.  
DR EMBL; BC011394; AAH11394.1; -.  
DR EMBL; BC022993; AAH22993.1; -.  
DR Genew; HGNC:10469; RTN3.  
DR MIM; 604249; -.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR InterPro; IPR003388; Reticulon.  
DR Pfam; PF02453; Reticulon; 1.  
DR PROSITE; PS50845; RETICULON; 1.  
KW Transmembrane; Endoplasmic reticulum.  
FT TRANSMEM 68 88 POTENTIAL.  
FT TRANSMEM 177 197 POTENTIAL.  
FT DOMAIN 48 236 RETICULON.  
SQ SEQUENCE 236 AA; 25609 MW; DDC6A4544ABCDFB7 CRC64;

Query Match 10.7%; Score 625.5; DB 1; Length 236;  
Best Local Similarity 59.1%; Pred. No. 1.7e-17;  
Matches 114; Conservative 41; Mismatches 37; Indels 1; Gaps 1;

Qy 972 SKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 1031  
| :| ||:||||:|||| ||| :| :|||| ||:| | :| : ||||| |  
Db 44 SSCAVHDLIFWRDVKKTGFEVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSV 103  
  
Qy 1032 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDLSLK 1091  
|||:||||:||||:|||| : :| | | | :| :| :| : |||||:|||||  
Db 104 IQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHINRALKLIIRLFLVEDLVDLSLK 163  
  
Qy 1092 FAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKI 1151  
|| ||: |||||:||||:||||| : :||:|:|:|:|: |||||:|:| | : ||  
Db 164 LAVFMWLMTYVGAVFNGITLLILAELLIFSVPVYKYKTQIDHYVGIARDQTKSIVEKI 223  
  
Qy 1152 QAKIPGL-KRKAD 1163  
|||:|:| :|:|:  
Db 224 QAKLPGIAKKKAE 236

#### RESULT 7

RTN3\_MOUSE

ID RTN3\_MOUSE STANDARD; PRT; 237 AA.

AC Q9ES97;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Reticulon protein 3.  
 GN RTN3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Huang X., Zhou Y., Qiang H., Yuan J., Qiang B.;  
 RT "Cloning and expression profile of a novel mouse cDNA encoding a human  
 RT RTN3 homolog.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum (Potential).  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF195940; AAG31360.1; -.  
 DR EMBL; BC014697; AAH14697.1; -.  
 DR MGD; MGI:1339970; Rtn3.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Transmembrane; Endoplasmic reticulum.  
 FT TRANSMEM 69 89 POTENTIAL.

FT TRANSMEM 167 187 POTENTIAL.  
FT DOMAIN 49 237 RETICULON.  
SQ SEQUENCE 237 AA; 25428 MW; EB60A0A7AC45F0DE CRC64;

Query Match 10.7%; Score 625.5; DB 1; Length 237;  
Best Local Similarity 59.1%; Pred. No. 1.7e-17;  
Matches 114; Conservative 41; Mismatches 37; Indels 1; Gaps 1;

Qy 972 SKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 1031  
| :| ||:||||:|||| ||| :| :|||| ||::|| :|: |||||:||||:| |  
Db 45 SSCAVHDLIFWRDVKKTGFFVGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRVYKSV 104  
  
Qy 1032 IQAIQKSDEGHPPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRFLVDDLVDLSLK 1091  
|||:||||:||||:||||: :: :| | | :|: ||| :| : ||||:|||||||  
Db 105 IQAVQKSEEGHPFKAYLDVDITLSSEAFHNYMNAAMVHVNKALKLIIRLFLVEDLVDLSLK 164  
  
Qy 1092 FAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDHYLGLANKSVKDAMAKI 1151  
|| ||: ||||:||||:|||||| : :||:|::|::| ||||:|:| | : ||  
Db 165 LAVFMWLMTYVGAVFNGITLLILAEELLVFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKI 224  
  
Qy 1152 QAKIPGL-KRKAD 1163  
|||:|: |::|  
Db 225 QAKLPGIAKKKAE 237

#### RESULT 8

##### RTN2\_MOUSE

ID RTN2\_MOUSE STANDARD; PRT; 471 AA.  
AC O70622; O70620;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like protein 1) (NSPLI).  
GN RTN2 OR NSPL1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.  
RC STRAIN=FVB/N, and 129/Sv; TISSUE=Cerebellum, and Skeletal muscle;  
RX MEDLINE=98191726; PubMed=9530622;  
RA Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;  
RT "Molecular cloning of a novel mouse gene with predominant muscle and  
RT neural expression."  
RL Mamm. Genome 9:274-282(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Retina;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum  
 CC (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=Brain;  
 CC IsoId=O70622-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Muscle;  
 CC IsoId=O70622-2; Sequence=VSP\_005650, VSP\_005651;  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in neural and muscular  
 CC tissues.  
 CC -!- SIMILARITY: Contains 1 reticulon domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF038537; AAC14906.1; -.  
 DR EMBL; AF038537; AAC14907.1; -.  
 DR EMBL; AF038538; AAC14908.1; -.  
 DR EMBL; AF038539; AAC14909.1; -.  
 DR EMBL; AF093624; AAD13195.1; -.  
 DR EMBL; BC031370; AAH31370.1; -.  
 DR MGD; MGI:107612; Rtn2.  
 DR InterPro; IPR003388; Reticulon.  
 DR Pfam; PF02453; Reticulon; 1.  
 DR PROSITE; PS50845; RETICULON; 1.  
 KW Endoplasmic reticulum; Alternative splicing; Transmembrane.  
 FT TRANSMEM 295 315 POTENTIAL.  
 FT DOMAIN 272 471 RETICULON.  
 FT VARSPLIC 1 267 Missing (in isoform 2).  
 FT /FTId=VSP\_005650.  
 FT VARSPLIC 268 271 PLLL -> MGSK (in isoform 2).  
 FT /FTId=VSP\_005651.  
 SQ SEQUENCE 471 AA; 51346 MW; 9BBD8F372CF63AD3 CRC64;

Query Match 8.7%; Score 509; DB 1; Length 471;  
 Best Local Similarity 28.6%; Pred. No. 1e-12;  
 Matches 146; Conservative 84; Mismatches 175; Indels 106; Gaps 13;

```

Qy      690  PDFSNYSEIAKFEKSVPEHAELVEDSSP--ESEPVDLFSDDSIPEVPQTQEEAVMLMKES 747
      |::: | |:: | | | | |::|
Db      6   PVFAHCKEAPSTASSTPDSTEGGNDDSDFRELHTAREFSED----- 46

Qy      748  LTEVSETVAQH-----KEERLSASPQELGKPYLESFQPNLHST 785
      | ||:| ::: |:: | | |
Db      47  --EEEETTSQDWGTPRELTFSYIAFDGVVSGGRRDSVVRPRPQGRSVSEPRDPQQSG 104

Qy      786  KDAASNDIPTLTKK-----EKISLQMEEFNTAIYSNDDLSSKEDK 826
      : ||::: |:: |::: : | | |
Db      105  LGDSLESIPSLSQSPPEGRRGDPDPVPPAERPLEELRLRLDQLGWVVR-----AGSGED- 159

Qy      827  IKESETFSDSSPIEIIDFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCL 886
      | |:::| :| | : | : | | |::: :| |
Db      160  ----SATSSSTPLE--NEEPDGLLEASE-----AGEETNLEL----RLAQ-----SL 195

Qy      887  ELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTK 946
      | :: : | : | | | : | | : :: : |
Db      196  HLQLEVLTPQLSPSSGTPQAHTPSPQRSQDSNS--GPDDEPLLNVVEEHWRLLLEQEPITA 253

Qy      947  EAEKKLPSDTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS 1006
      : |::: | :| | | | | : : | | : | | |
Db      254  QC----LDSTDQSEFMLEPLL-----LVADLLYWKDTRTSGAVFTGLMASLLCLLHFS 302

Qy      1007  IVSVTAYIALALLSVTISFRIYKGVIAIQKSDGHPFRAYLESEVAISEELVQKYSNSA 1066
      | | | | |::| | | | |:: |::: : | | | | |::: : | : |
Db      303  IVSVA AHLALLGLCATISLRVYRKVLQAVHRGDTNPFQAYLDMDLTLTREQTERLSQQI 362

Qy      1067  LGHVNSTIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIY 1126
      | | | | : | | | | | | | | | : : | : | | | | | : : | : : |
Db      363  ASHVVSTATQLRHFFLVEDLVDLSLKLALLFYILTFVGAIFNGLTLVILGVVALFTVPLLY 422

Qy      1127  ERHQVQIDHYLGLANKSVKDAMAKIQAKIPG 1157
      : | | | | : | : | | | | |
Db      423  ROHOAOIDQYVGLVTNOLSHIKAKIRAKIPG 453

```

## RESULT 9

RTN2 HUMAN

```

RT gene, a member of a gene family encoding reticulons."
RL Genomics 51:98-106(1998).
RN [2]
RP SEQUENCE OF 108-545 FROM N.A. (ISOFORM RTN2-B).
RC TISSUE=Brain;
RX MEDLINE=98191726; PubMed=9530622;
RA Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT "Molecular cloning of a novel mouse gene with predominant muscle and
RT neural expression.";
RL Mamm. Genome 9:274-282(1998).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=RTN2-A;
CC IsoId=O75298-1; Sequence=Displayed;
CC Note=Isoform RTN2-C is produced by alternative initiation at
CC Met-341 of isoform RTN2-A;
CC Name=RTN2-B;
CC IsoId=O75298-2; Sequence=VSP_005649;
CC Event=Alternative initiation;
CC Comment=2 isoforms, RTN2-A (shown here) and RTN2-C, are produced
CC by alternative initiation at Met-1 and Met-341;
CC -!- TISSUE SPECIFICITY: ISOFORM RTN2-C IS HIGHLY EXPRESSED IN SKELETAL
CC MUSCLE.
CC -!- SIMILARITY: Contains 1 reticulon domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF004222; AAC32542.1; -.
DR EMBL; AF004223; AAC32543.1; -.
DR EMBL; AF004224; AAC32544.1; -.
DR EMBL; AF038540; AAC14910.1; -.
DR Genew; HGNC:10468; RTN2.
DR MIM; 603183; -.
DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.
DR GO; GO:0004871; F:signal transducer activity; NAS.
DR GO; GO:0007165; P:signal transduction; NAS.
DR InterPro; IPR003388; Reticulon.
DR Pfam; PF02453; Reticulon; 1.
DR PROSITE; PS50845; RETICULON; 1.
KW Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW Alternative initiation.
FT CHAIN 1 545 RETICULON PROTEIN 2, ISOFORM RTN2-A.
FT CHAIN 341 545 RETICULON PROTEIN 2, ISOFORM RTN2-C.
FT INIT_MET 341 341 FOR ISOFORM RTN2-C.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
FT DOMAIN 345 545 RETICULON.
FT VARSPLIC 272 344 Missing (in isoform RTN2-B).
FT /FTId=VSP_005649.

```

SQ SEQUENCE 545 AA; 59263 MW; 971FD2F909E1E9E6 CRC64;

Query Match 8.4%; Score 490; DB 1; Length 545;  
Best Local Similarity 28.3%; Pred. No. 6.3e-12;  
Matches 160; Conservative 76; Mismatches 202; Indels 128; Gaps 17;

```
Qy      668 EAPYISIACDLIKETKLSTEPSPDFSNYSEI---AKFEKSVPEHAELVEDSSPESEPVDL 724
      |||      | ||| | |:: |:      :| : |      : :|
Db      13 EAP-----STASSTPDSTEGGNDDSDFRELHTAREFSEEDDEEETTSQDWGTPRELTFSY 66

Qy      725 FSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELGKPYL-ESFQPNLH 783
      : | :      ::      : ||| || : | | : : | | :| :|
Db      67 IAFDGVVGSGGRDSTARRPRPQGRSVSEPRDQHPQPSLGDLSLESI--PSLSQSPEPGRR 124

Qy      784 STKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDDSSPIEIID 843
      | |      |:      | : |:::      | ||      : | |::|
Db      125 GDPDTAP---PSERPLEDLRLRLDHLGWVARGT----GSGED-----SSTSSSTPLE--- 169

Qy      844 EFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKNIYPKDEV 903
      | : |      :| || : |      || || : |
Db      170 -----DEEP---QEPNRLETGEAGE-----EL--DLRLRLAQPSSE 201

Qy      904 HVSDEFSENRSSVSKASISP-----SNVSALEPQTE----- 934
      :: : |      : : || || || |
Db      202 VLTPQLSPGSGTPQAGTPSPSRSRDSNSGPPEEPLLEEEKQWGPLEREFVVRGQCILDSTDQ 261

Qy      935 -----MGSIVK--SKSLTKEAEKKLP-----S 954
      :| : : || | :|
Db      262 LEFTVEPRLLGTAMEWLKTSLLLAVYKTVPILELSPPLWTAIGWVQRGPTPPTPVLRVLL 321

Qy      955 DTEKEDRSLSAVLSAELSK---TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVT 1011
      | || | | | | : | ||||:| : :||| : || | |||||
Db      322 KWAKSPRS-SGVPSLSLGDADMGSKVADLLYWKDTRTSGVVFTGLMVSLLCLLHFSIVSVA 380

Qy      1012 AYIALALLSVTISFRIYKGVIAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVN 1071
      |::|| || ||| |::: |::|: : | :||:||||: :: :| : :| :
Db      381 AHLALLLLCGTISLRVYRKVLQAVHRGDGANPFQAYLDVDLTLTREQTERLSHQITSRVV 440

Qy      1072 STIKELRRLFLVDDLVDLSLKFAVLMWVFTYVGFALFNGLTLLILALISLFSIPVIYERHQV 1131
      | :|| |||:|||||| |:: :| :||:||||||| :| ||:|::| :||
Db      441 SAATQLRHFFLVEDLVDLSLKLALLFYILTFVGAIENGLTLLILGVIGLFTIPLLYRQHQA 500

Qy      1132 QIDHYLGLANKSVKDAMAKIQAKIPG 1157
      ||| |::| : |||:|||||
Db      501 QIDQYVGLVTNQLSHIKAKIRAKIPG 526
```

RESULT 10

PCLO\_HUMAN

ID PCLO\_HUMAN STANDARD; PRT; 5147 AA.  
AC Q9Y6V0; O43373; O60305; Q9BVC8; Q9UIV2; Q9Y6U9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Piccolo protein (Aczonin) (Fragments).  
GN PCLO OR ACZ OR KIAA0559.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 1-759 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99439764; PubMed=10508862;  
 RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,  
 RA Kilimann M.W.;  
 RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active  
 RT zones, shares homology regions with rim and bassoon and binds  
 RT profilin.";  
 RL J. Cell Biol. 147:151-162(1999).  
 RN [2]  
 RP SEQUENCE OF 552-4404 FROM N.A.  
 RA Kraemer J., Wollam C., Wohldmann P., McGrane B.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:31-39(1998).  
 RN [4]  
 RP SEQUENCE OF 4405-4439 FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 4405-5147 FROM N.A.  
 RA Kalicki J., Elliott G.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle



```

CC      trafficking (By similarity).
CC      -!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin (By
CC      similarity).
CC      -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC      synaptic junctions (By similarity).
CC      -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Comment=Additional isoforms seem to exist;
CC      Name=1;
CC      IsoId=Q9Y6V0-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q9Y6V0-2; Sequence=VSP_003923, VSP_003924, VSP_003925,
CC      VSP_003926, VSP_003927;
CC      Note=No experimental confirmation available;
CC      -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC      phospholipids. Calcium binds with low affinity but with high
CC      specificity and induces a large conformational change.
CC      -!- SIMILARITY: Contains 2 C2 domains.
CC      -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y19188; CAB60727.1; -.
DR      EMBL; AC004903; AAD20936.1; -.
DR      EMBL; AC004886; AAD21789.1; -.
DR      EMBL; AB011131; BAA25485.1; -.
DR      EMBL; BC001304; AAH01304.1; -.
DR      EMBL; AC004082; AAB97937.1; -.
DR      PIR; T00634; T00634.
DR      HSSP; P04410; 1A25.
DR      Genew; HGNC:13406; PCLO.
DR      MIM; 604918; -.
DR      GO; GO:0005856; C:cytoskeleton; NAS.
DR      GO; GO:0045202; C:synaptic junction; ISS.
DR      GO; GO:0005509; F:calcium ion binding; ISS.
DR      GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
DR      GO; GO:0005522; F:profilin binding; ISS.
DR      GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
DR      GO; GO:0016080; P:synaptic vesicle targeting; ISS.
DR      InterPro; IPR000008; C2.
DR      InterPro; IPR001565; Synaptotagmin.
DR      PRINTS; PR00360; C2DOMAIN.
DR      PRINTS; PR00399; SYNAPTOTAGMN.
DR      SMART; SM00239; C2; 2.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS50004; C2_DOMAIN_2; 2.
KW      Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
KW      Repeat; Alternative splicing.
FT      NON_TER      1      1
FT      DOMAIN      400      465      10 X 10 AA TANDEM APPROXIMATE REPEATS OF
FT      P-A-K-P-Q-P-Q-Q-P-X.

```

FT	ZN_FING	499	523	C4-TYPE (POTENTIAL).
FT	ZN_FING	969	992	C4-TYPE (POTENTIAL).
FT	NON_CONS	1010	1011	
FT	DOMAIN	2300	2325	POLY-PRO.
FT	DOMAIN	4391	4442	PDZ.
FT	DOMAIN	4544	4633	C2 DOMAIN 1.
FT	DOMAIN	5031	5121	C2 DOMAIN 2.
FT	VARSPPLIC	4404	4404	S -> SGNGLGIRIVGGKEIPGHSGEIGAYIAKILPGGSAE
FT				QTGKLMEG (in isoform 2).
FT				/FTId=VSP_003923.
FT	VARSPPLIC	4534	4534	K -> KPTDGTKVVSHPITGEIQ (in isoform 2).
FT				/FTId=VSP_003924.
FT	VARSPPLIC	4576	4576	G -> GQVMVQNAS (in isoform 2).
FT				/FTId=VSP_003925.
FT	VARSPPLIC	4757	4761	TAHKS -> SKRRK (in isoform 2).
FT				/FTId=VSP_003926.
FT	VARSPPLIC	4762	5147	Missing (in isoform 2).
FT				/FTId=VSP_003927.
SQ	SEQUENCE	5147 AA;	563537 MW;	CD5D84990498CD3C CRC64;

Query Match 5.3%; Score 308.5; DB 1; Length 5147;  
 Best Local Similarity 21.4%; Pred. No. 0.00076;  
 Matches 264; Conservative 150; Mismatches 438; Indels 383; Gaps 63;

Qy	3	DIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERKPA	62
		:        :    : :   : :   :	
Db	251	DIVRGESVKPSLPSPSKPP-----IQQPTPGKPPAQQPGHEKSQPG-----PAKPP	296
Qy	63	AGLSAAAVP-----PAAAAPLLDFSSDSVPPA-PRGPLPAAPPA-----A	101
		: :	
Db	297	AQPSGLTKPLAQQPGTVKPPVQPPGTTKPPAQPLG--PAKPPAQQTGSEKPSSEQPGPKA	354
Qy	102	PERQPSWERSPA-APAPSLP-----PAAAVLPSKLPEDDEPPA---	138
		:   :        :          :   :	
Db	355	LAQPPGVGKTPAQQPGPAKPPTQQVGTGPKPLAQQPLQSPAKAPGPTKTTPAQTKPPSQQP	414
Qy	139	---RPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDETLFALPAASEFVIPSSAEKIM	195
		:            : :        :     :	
Db	415	GSTKPPPPQQPGPAKP--SPQQPGSTKPPSQPGS-----AKPSAQQ--PSPAKPSA	461
Qy	196	DLMEQPGNTVSSSGQEDFPSVLLETAASLPS-----LSPL-STVSFKEH-GYLGNL SAVS	247
		:   :   :   : :     :    :   : :	
Db	462	QQFTKPVSTGFGKPLQPPPTVSPSAKQPPSQGLPKTICPLCNTTELLLVPEKANFNTCT	521
Qy	248	SSEGTI-----EETLNEASKEL-----PERATNPFVNRDL	277
		:   :     :     :	
Db	522	ECQTTVCSLCGFNPNPHLTEAKEWLCINCMKRALGGDLAPVPSSPQPKLKTAPVTTTSA	581
Qy	278	AEFSELEYSEMGSFSGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGK	337
		: : :    : :     :   :   :	
Db	582	VSKSSPQPQQTSPKKDAAPKQDLKAPEPKPPPLVKQPTLHGSPSAKAKQPPEADSLSK	641
Qy	338	EDRVVSPEKTMDIFNEMQMSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAAARANVES	397
		:     : :            :     :   : :	
Db	642	P----APPKEPSVPSEQDK---APV----ADDKP-KQPKMVKPTTD---LVSSSSATTKP	686
Qy	398	KVDRKCLEDLSLEQKS---LGKDSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTT	454

```

      :  ::  |::  |  ||  ::  |||  |  |  |  |
Db      687 DIPSSKVQSQAEKTTPLKTD SAKPSQ--SFPPTGEKV-----TPFDSKAI PRP 734

Qy      455 ANTFPLLE--DHTSENKTD-----EKKIEERKAQIITEKTS PK----- 490
      |:  ::  :||:|  :||  |  :|||  |  |||
Db      735 ASDSKIISHPGPSSES KGQKQVDPVQKKEEPKKAQ---TKMSPKPD AKPMPKGSPTPPGP 791

Qy      491 --TSNPFLVAVQDS----EADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEAT 544
      |:  :  |  |  |  :  |  :||  |  |  |  |
Db      792 RPTAGQTVPTPQQSPKPQEQRRLSLNLGSITDAPKSQ-----PTTPQET-----VT 838

Qy      545 GTKIA YETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSA 604
      |  :  :  |  |  |  :  |  |  :  |  :  ||  |  |::
Db      839 GKLF GFGASI-FSQASNLISTAGQPG----PHSQSGPGAP-----MKQAPAPSQPPTS 886

Qy      605 GASVVQPSVSPLEAPP-----PVS YD-----SIKLEP-----ENPPPYEE 639
      |  |  :|||  ||  :  :  |||  :  |||  ::
Db      887 QG----PPKSTGQAPPAPAKSIPVKKETKAPAAE KLEPKAEQAPT VKRTETETKPPPIKD 942

Qy      640 AMNVALK----ALGTK-----EGIKEPESFNAAVQ-----E 666
      :  ::  :  |  ||  |  ||:|  :||  :
Db      943 SKSLTAEPQKAVLPTKLEKSPKPESTCPLCKTELNIGSKDPPNFNTCTECKNQVCNLCGF 1002

Qy      667 TEAPYIS IAC-----DLIK-----ETKLSTEPSPDFSNYSEIAKFEKSVPEHA 709
      |:::  |  |  :  |  |  |  |  |  :  :  :  :||
Db      1003 NPTPHLTENCQTQRAISGQLDIRKMPPAPSGPKASEMPVP-----TESSSQKTAVPPQV 1057

Qy      710 ELV----EDSSPESEPV-----DLFSDDSIPEV---PQTQEEAVMLMKESL TEVSETVA 756
      :||  ::  |::|  :  |  :  ||  :  |  |||:  |  |  :  :  |
Db      1058 KLVKKQE QEVKTEAEKVILEKVKETLSMEKIPPMVTTDQKQES-KLEKDKASALQEKKP 1116

Qy      757 QHKEERLSAS PQEL----GKPYLESFQPNLHSTKDAASNDIPTLT KKEKISLQMEEFNTA 812
      :||:|  ::  ||||  :|  |  |  :  |  |  |
Db      1117 LP EEKKLIPEEEKIRSEEKKPLLEEKKPTPEDKK-----LLPEAKTSAPEEQ---- 1163

Qy      813 IYSNDDLSSK----EDKI-----KESETFSDSSPIEIIDEFPTFVSAKDDS- 855
      |||  |:  |::  |  :|  :  |  :  |  |  :  :||
Db      1164 ---KHDLLKSQVQIAEEKLEGRVAPKTVQEGKQPQTKMEGLP----SGTPQSLPKEDDKT 1216

Qy      856 -----PKLAKEYTDLEVSDKSEIANIQ-----SGADSLPCLELPCD 891
      |  ||  :  |  |||:  ::  |  |  |  |  |
Db      1217 TKTIKEQPQPPCTAKPDQEKE-DDKSDTSSSQPKSPQGLSDTGYSSDGISSSLGEIP-- 1273

Qy      892 LSFKN IYPKDEVHV-----SDEFSENRSSVSK---ASISPSNVSALEPQTEMGSIVKSKS 943
      ::  |  ||  :  |  ||:  |  |  |  :  :  :  |  ||  |  :  ||:
Db      1274 ----SLIPTDEKDILKGLKKDSFSQESSPSSPSDLAKLESTVLSILEAQASTLADEKSEK 1329

Qy      944 LTK--EAEKKLPSDTEKEDRSLSAVLSAELSKTSV 976
      |:  |  :  |  ||  |  |||  |  :|:  :
Db      1330 KTQPHEVSPEQPKDQEK-TQSLSETLEITISEEEI 1363

```

RESULT 11

PCLO\_CHICK

ID PCLO\_CHICK STANDARD; PRT; 5120 AA.

AC Q9PU36;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Piccolo protein (Aczonin) (Fragment).  
 GN PCLO OR ACZ.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99439764; PubMed=10508862;  
 RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,  
 RA Kilimann M.W.;  
 RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active  
 RT zones, shares homology regions with rim and bassoon and binds  
 RT profilin.";  
 RL J. Cell Biol. 147:151-162(1999).  
 CC -!- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle  
 CC trafficking (By similarity).  
 CC -!- SUBUNIT: Interacts with Rabac1/Pral and profilin (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of  
 CC synaptic junctions (By similarity).  
 CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and  
 CC phospholipids. Calcium binds with low affinity but with high  
 CC specificity and induces a large conformational change.  
 CC -!- SIMILARITY: Contains 2 C2 domains.  
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y19187; CAB60725.1; -.  
 DR HSSP; P04410; 1A25.  
 DR GO; GO:0045202; C:synaptic junction; ISS.  
 DR GO; GO:0005509; F:calcium ion binding; ISS.  
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.  
 DR GO; GO:0005522; F:profilin binding; ISS.  
 DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.  
 DR GO; GO:0016080; P:synaptic vesicle targeting; ISS.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR001478; PDZ.  
 DR InterPro; IPR001565; Synaptotagmin.  
 DR InterPro; IPR008899; Znf\_piccolo.  
 DR Pfam; PF00168; C2; 2.  
 DR Pfam; PF00595; PDZ; 1.  
 DR Pfam; PF05715; Zf\_piccolo; 2.  
 DR PRINTS; PR00399; SYNAPTOTAGMN.  
 DR SMART; SM00239; C2; 2.  
 DR SMART; SM00228; PDZ; 1.

DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 2.  
 DR PROSITE; PS50106; PDZ; 1.  
 KW Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;  
 KW Repeat.  
 FT NON\_TER 1 1  
 FT DOMAIN 258 357 10 X 10 AA TANDEM APPROXIMATE REPEATS OF  
 FT P-A-K-P-Q-P-Q-Q-P-X.  
 FT ZN\_FING 368 392 C4-TYPE (POTENTIAL).  
 FT ZN\_FING 836 859 C4-TYPE (POTENTIAL).  
 FT DOMAIN 2324 2343 POLY-PRO.  
 FT DOMAIN 4414 4493 PDZ.  
 FT DOMAIN 4627 4726 C2 DOMAIN 1.  
 FT DOMAIN 5003 5094 C2 DOMAIN 2.  
 SQ SEQUENCE 5120 AA; 560751 MW; A658D9891B65B412 CRC64;

Query Match 5.1%; Score 300; DB 1; Length 5120;  
 Best Local Similarity 20.4%; Pred. No. 0.0016;  
 Matches 258; Conservative 152; Mismatches 464; Indels 388; Gaps 56;

Qy 14 TDSPPRPPPAFKYQFVTEPEDEEDEEE-----EEDDEEDDEDLEEELEVLERKPAAGLSAA 68  
 :|| | : | ||: | : ||: :: : :: : : | | :  
 Db 90 SDSDAHEEAGRKQKVTQKEQGKPEEQRLAKHPSQQQSPKLVQQQGPVKPTPQQTESSK 149  
 Qy 69 AVP-----PAAAAPLLDFFSSDSV--PPAPRGP----LPAAPPAAPERQPSWE 109  
 || | : | | : : | || || | : | | : | :  
 Db 150 PVPQQQQPGEPKQGQKPGPSHP-GDSKAEQVKQPPQPRGFPQKSQLOQQSEPTKPGQQQTSA 208  
 Qy 110 RSPAAPAPSLP--PAAAVLPSKLPEDDEP-----PAR-----P 140  
 : : | | | : | : | : | || | | |  
 Db 209 KTSAGPTKPLPQQPDSAKTSSQAPPPTKPSLQQSGSVKQPSQQPARQGGPVKPSAQQAGP 268  
 Qy 141 PPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDET--LFALPAASEPVIPISS--AEKIMD 196  
 | | | : | | | : : | : : | : | || |  
 Db 269 PKQQPGSEKPTAQQTGPAKQPPQP-GPGKTPLQQTGPVKQVPPQAGPTKPSQTAGAAS 327  
 Qy 197 LMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYLGNL SAVSSSEG TIEET 256  
 | : || | | : | : : | | | : |  
 Db 328 LAQQPGLTKPPGQQPGPEKPLQQQA-----STTQPVESTPKKT 366  
 Qy 257 L-----NEASKELPERAT-----NPFVNRDLAEFSE-----LEYS 286  
 | ||: | || | : | | : :  
 Db 367 FCPLCTTTTELLLHTPEKANYNTCTQCHTVVCSLCGFNP--NPHITEIKEWLCLNCQMORA 424  
 Qy 287 EMGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPE- 345  
 | | | : | | | | | : | :  
 Db 425 LGGDLASGHGPGPQ-----LPPPKQKTPTPASTAKPSPQLQPGQKKDASPKPDP 473  
 Qy 346 -----KTMDIFNEMQMSVVAPVREEYADFKEPFEQAWEVKDTYEGSRDVLAAARANVESK 398  
 | : : | ||: : : : : | | : |  
 Db 474 SQQADSKKVPVQKKQPSMPGSPVVKSKQTHAEPSTGQQI-DSTPKSDQVKPTQA----- 527  
 Qy 399 VDRKCLED SLEQKSLGKD-----SEGRNEDASFPSTPEPVKDSSRAYITCASFTSA 449  
 | : | : | : | : | : | : : :  
 Db 528 -----EEKQNQPSIQKPTMDTVPTSAAPGVKQDLADPQSP-----STQQKVTDSPMPET 576  
 Qy 450 TESTTANTFPLLEDHTSENKTDEKKIEERKAQIITEKTSP-----KTSNPFLV 497

Db	577	TK-----PPADTHPAGDKPDSKPL-----PQVSRQKSDPKLASQSGAKSDAKTQKPSEP	625
Qy	498	A-VQDSEADYVT-----TDTLSKVTEAAVSNMPE-----	525
Db	626	APVKDDPKKLQTKPAPKPDTPAPKGPQAGTGPRPTSAQPAPQPPQKTPEQSRRFSLN	685
Qy	526	--GLT-----PDLVQEACESEL-----NEA-----TGTKIAYETKVDLVQTSEAIQ	564
Db	686	LGGITDAPKPQPTTPQETVTGKLFGFGASIFSQASSLISTAGQPGSQTSGPAPPATKQPQ	745
Qy	565	ESLYPTAQLCPSFEEAEATPSP-VLPDIVMEAPLNS--LLPSAGASVVQPSVSPLEAPP	621
Db	746	PPSQPPASQAPPKEAAQAQPPPKAAPTKKETKPLASEKLGPMASDSTLTTKGSDLEKKPS	805
Qy	622	VSYDSIKLEPENPPPYE-----EAMNVALKALGTKEGIKEPESFNAA-----	663
Db	806	LAKDSKHQTAEAKKPAELSEQEKASQPKVSCPLCKTGLNIGSKDPPNFNTCTECKKVVCN	865
Qy	664	-----VQETEAPYISIA-----DLIK-----	680
Db	866	LCGFNPMPHIVEVQE----WLCLNCQTQRAMSGQLGDMGKVPLPKLGPSQPVSKPPATPQ	921
Qy	681	-----ETKLSTEPSPDFSNYSEIAKFEKSV--EHAEL-----VEDSS	716
Db	922	KQPVPVAVSHSPQKSSTPPTPAATKPKKEEPSVPKEVPKLQQKGLEKTLKADKIQQGIQKED	981
Qy	717	PESEPVDLF---SDDSIPEVPQTQEEAVMLMKESLTVSETVAQH-KEERLSASPQELGK	772
Db	982	AKSKQGKLFKTPSADKIQRVSQKEDSRLQQTKLTKTPSSDKILHGVQKEDIKFQEAQLAK	1041
Qy	773	-----PYLESFQPNLHSTKDA---ASNDIPTLTKEKISLQMEEFNTAIYSN-----	816
Db	1042	IPSADKILHRLQKEDPKLQQMMAKALSADKIQPEAQKEDVQLQEVRLSKAVSADKIQHG	1101
Qy	817	--DDL-----LSSKEDKI----KESETFSDSPIEII-DEFPTFVSA-----	851
Db	1102	IQKDLNLQHVKIEKTSSVEKIQEAQKESKLQQDKLPKTLSEDKIPATVSSDHKKLLSKSE	1161
Qy	852	KDDSPKLAKEYTDLEVSDKSEIANIQSG--ADSLPCLELPCDLSFKNIYPKDEVHVSDE-	908
Db	1162	EDKKPELLEKSTPHPKDKKEQITAETTGHITEQKVEVEAPCD----KLHEKKQEDVKKED	1217
Qy	909	----FSENRSSVSKA--SISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPDTEKEDRS	962
Db	1218	LTTGIPQMVSKPEKAEEEKTPVPVSRL-PRSDHVEAVREK-IEKEDDK---SDTSSSQQQ	1272
Qy	963	LS 964	
Db	1273	KS 1274	

RESULT 12

CPN\_DROME

ID CPN\_DROME STANDARD; PRT; 865 AA.

AC Q02910;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Calphotin.  
 GN CPN OR CAP.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Canton-S;  
 RX MEDLINE=93165729; PubMed=8094559;  
 RA Martin J.H., Benzer S., Rudnicka M., Miller C.A.;  
 RT "Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Canton-S;  
 RX MEDLINE=93165730; PubMed=8434015;  
 RA Ballinger D.G., Xue N., Harshman K.D.;  
 RT "A Drosophila photoreceptor cell-specific protein, calphotin, binds  
 calcium and contains a leucine zipper.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).  
 CC -!- FUNCTION: Might function as a calcium-sequestering "sponge" to  
 CC regulate the amount of free cytoplasmic calcium. It binds 0.3 mole  
 CC of Ca(2+) per mole of protein.  
 CC -!- SUBUNIT: Homodimer (Probable).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; hypodense compartment.  
 CC -!- TISSUE SPECIFICITY: Soma and axons of photoreceptor cells of  
 CC compound eyes and ocelli.  
 CC -!- DEVELOPMENTAL STAGE: Expressed early in photoreceptor cell  
 CC development.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L02111; AAA28405.1; -.  
 DR EMBL; L05080; AAA28420.1; -.  
 DR PIR; A47282; A47282.  
 DR PIR; A47283; A47283.  
 DR FlyBase; FBgn0010218; Cpn.  
 DR GO; GO:0005509; F:calcium ion binding; IDA.  
 KW Calcium-binding.  
 FT CONFLICT 36 36 A -> AVAPAVVA (IN REF. 2).  
 FT CONFLICT 43 43 I -> T (IN REF. 2).  
 FT CONFLICT 64 64 I -> V (IN REF. 2).  
 FT CONFLICT 76 76 T -> A (IN REF. 2).  
 FT CONFLICT 100 100 P -> PP (IN REF. 2).  
 FT CONFLICT 126 127 VQ -> AP (IN REF. 2).  
 FT CONFLICT 154 154 I -> V (IN REF. 2).  
 FT CONFLICT 160 160 S -> T (IN REF. 2).

FT CONFLICT 534 534 A -> E (IN REF. 2).  
 FT CONFLICT 699 699 I -> T (IN REF. 2).  
 FT CONFLICT 703 703 V -> L (IN REF. 2).  
 FT CONFLICT 721 721 D -> E (IN REF. 2).  
 SQ SEQUENCE 865 AA; 84781 MW; 2110417E0B0E7CFE CRC64;

Query Match 5.1%; Score 295.5; DB 1; Length 865;  
 Best Local Similarity 21.3%; Pred. No. 0.00024;  
 Matches 217; Conservative 127; Mismatches 379; Indels 295; Gaps 41;

Qy 62 AAGLSAAAVPPAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPER----QPSWERSPAAPAP 117  
 :| ::| | | | | :| | | | :| | | | | :| | | |  
 Db 11 SAPVAAPVTPSAVAAPVQVVSAPAAVAPAPAAPIAVTPVAPPPTLASVQPATVTIP-APAP 69  
  
 Qy 118 ----SLPPAAVLPSKLPEDDEPPARPPPPPPAGA-----SPLAEPAPPSTP----- 161  
 | : | | : | : | | | | | | | : | : | | | |  
 Db 70 IAAASVTPVASVAPPVVAAPTTPAASPSTPVAVAQIPVAVSAPVAPPVVAATPTPVVQIP 129  
  
 Qy 162 -AAPKRRGSGSVDETLFALP--AASEP----VIPSSAEKIMDLMEQPGNT---VSSGQED 211  
 | | | : | | | | | | | : | | | | : :  
 Db 130 VAAP-----VIATPPVAASAPTPAAVTPVISPVIASPPVVPANTTVPVAAPVAA 178  
  
 Qy 212 FPSVLLETAASL-PSLSPLSTVSFKEHGYLGNLSAVSSSEGTEETLNEASKELPERATN 270  
 | : : | | | : : | | : : : | | | : | : | |  
 Db 179 VPAAVPVVAPVLAPAVAPAVAPVVAETPAPPPVAEIPVA--TIPECVAPLIPEVSVVATK 236  
  
 Qy 271 PFVNRDLAEFSELEYSEMGSSFKGSPKGESAILVENTKEEVIVRSKDKEDLVCSAALHSP 330  
 | | | | | | | | | | | | | | | | | | | |  
 Db 237 PLA-----AAEPVVVAPPATETFPVVAAPAAASP 263  
  
 Qy 331 QESPVGKEDRVVSPEKTMDIFNEMQMSVVAPVREEYADFKPFEQAWVEKDTYEGSRDVLA 390  
 | | : | : : : | | | | | | | | | | | | | |  
 Db 264 HVS-----VAP-----AVETAVVAPV----- 279  
  
 Qy 391 ARANVESKVDRKCLEDSELEQKSL----GKDSEGRNEDASFPSTPEPVKDSSRAYITCASF 446  
 | : | | | : | : | : | : | | | | : : : :  
 Db 280 -SASTEPPVAAATLTTPAPETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAETP 338  
  
 Qy 447 TSATESTTANTFPLEDHTSENKTDEKKIEERKAQIITEKTSPKTSNPFVLVAVQDSEADY 506  
 | : : | | : : : | | | | | | | | | | : | : |  
 Db 339 EVASVAVAETTPPVVPPVAAES-----IPAPVVATTPVPATLAVTDPD--- 381  
  
 Qy 507 VTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQES 566  
 | | : | | : | : | | | | | | | | | | | | | |  
 Db 382 -----VTASAVPELPPVIAPSPVPSA-----VAETPVDLA-----PPV 414  
  
 Qy 567 LYP-TAQLCPSF--EEAEATPSPVLPDIVMEAPLNSLLPSAGASVVQPSVSPLAEP---- 619  
 | | | : | : | | | | : : | | : | : | | | | |  
 Db 415 LPPVAAEPVPAVVAEETPETPAPASAPVTI-AALD--IPEVAPVIAAPSDAPAEAPSAAA 471  
  
 Qy 620 PPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFNAAVQETEAPYISIIACD-- 677  
 | | | | | | | | | | | | | | | | : : :  
 Db 472 PIVS-----TPPT-----TASVPETTAPPAAVPTEPI 498  
  
 Qy 678 ---LIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAE--LVEDSSPESEPVDLFS-DSIP 731  
 : : | : | : | : : | : : | | : : | : | | : |  
 Db 499 DVSVLSEAAIETPVAPPVEVTTEVAVADVAPPEAAADLIIEPVEPPAPIPDLLQTTSPV 558



QY	732	EVPTQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELGKPYLESFQPNLHSTKDAASN	791
		: : : :    :      : : :   :   :   : :	
Db	559	AVEAAESTSSSPIPETSLPPPNEAVA---SPEVAVAPITAPEPIPEP-EPSLATPTEPIPV	614
QY	792	DIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIEIIIDEFPTFVSA	851
		:   : : :   : :   : :   :         :	
Db	615	EAPV-----VIQEAVDAVEVPVTETSTSIPIETTVEFPEAVAE	651
QY	852	KDDSPKLAKEYTDLEV-SDKSEIANIQSGAD----SLPCLEL-----PCDLSFKNIYP-	899
		: :   : : : :    :   : :   : :   :	
Db	652	KVLDPAI---TEAPVTTQEPDVANINDGAPATEITTPAVEIVTAAAEVSDIAIPVIDPP	707
QY	900	-KDEVHVSDEFSENRRSSVSKASISPSNVSALEPQTEMG-----	936
		:   :   : : : :   :   : :	
Db	708	VPQEIAVA-EIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAAEVPITAGDN	766
QY	937	----SIVKSKSLTKEAEKKLPDSTEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGV	990
		:   : :     :   :   : :       :   :	
Db	767	PDNTSVGISEVVPTIAEKPVEEVPTSEIPEQSSSPSDSVPVAKITPLL--RDLQTTDV	822

## MAPB RAT

```

ID      MAPB_RAT      STANDARD;      PRT;      2459 AA.
AC      P15205; Q62958; Q9ER21; Q9QW92;
DT      01-APR-1990 (Rel. 14, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Microtubule-associated protein 1B (MAP 1B) (Neuraxin) [Contains: MAP1
DE      light chain LC1].
GN      MAP1B.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE OF 1-142 FROM N.A.
RC      STRAIN=Sprague-Dawley; TISSUE=Testis;
RX      MEDLINE=96257242; PubMed=8666295;
RA      Liu D., Fischer I.;
RT      "Isolation and sequencing of the 5' end of the rat microtubule-
RT      associated protein (MAP1B)-encoding cDNA.";
RL      Gene 172:307-308(1996).
RN      [2]
RP      SEQUENCE OF 96-2459 FROM N.A., DOMAIN, AND INDUCTION.
RC      STRAIN=Sprague-Dawley; TISSUE=Brain, and Glial tumor;
RX      MEDLINE=92347374; PubMed=1639092;
RA      Zauner W., Kratz J., Staunton J., Feick P., Wiche G.;
RT      "Identification of two distinct microtubule binding domains on
RT      recombinant rat MAP 1B.";
RL      Eur. J. Cell Biol. 57:66-74(1992).
RN      [3]
RP      SEQUENCE OF 1541-2459 FROM N.A., AND TISSUE SPECIFICITY.
RC      TISSUE=Spinal cord;
RX      MEDLINE=90059871; PubMed=2555150;
RA      Rienitz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,

```

RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;  
 RT "Neuraxin, a novel putative structural protein of the rat central  
 RT nervous system that is immunologically related to microtubule-  
 RT associated protein 5.";  
 RL EMBO J. 8:2879-2888(1989).  
 RN [4]  
 RP DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.  
 RX MEDLINE=97405699; PubMed=9260743;  
 RA Ma D., Nothias F., Boyne L.J., Fischer I.;  
 RT "Differential regulation of microtubule-associated protein 1B (MAP1B)  
 RT in rat CNS and PNS during development.";  
 RL J. Neurosci. Res. 49:319-332(1997).  
 CC -!- FUNCTION: The function of brain MAPS is essentially unknown.  
 CC Phosphorylated MAP1B may play a role in the cytoskeletal changes  
 CC that accompany neurite extension. Possibly MAP1B Binds to at least  
 CC two tubulin subunits in the polymer, and this bridging of subunits  
 CC might be involved in nucleating microtubule polymerization and in  
 CC stabilizing microtubules.  
 CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate  
 CC with MAP1A and MAP1B proteins.  
 CC -!- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,  
 CC cerebellum and cerebrum). Not expressed in liver, spleen, kidney,  
 CC heart or muscle.  
 CC -!- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic  
 CC nerve levels are high early in development but decrease during  
 CC postnatal development and are low in adults. In dorsal root  
 CC ganglia levels remain high throughout development.  
 CC -!- INDUCTION: By nerve growth factor.  
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC responsible for the binding of MAP1B to microtubules.  
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated  
 CC from MAP1B by proteolytic processing. It is free to associate with  
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region  
 CC of MAP1B (By similarity).  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: TO MAP1A.  
 CC -!- CAUTION: A C-terminal fragment of this protein (residues 1597 to  
 CC 2459) was originally described as neuraxin in Ref.3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U52950; AAB17068.1; -.  
 DR EMBL; X60370; CAC16162.1; -.  
 DR EMBL; X16623; CAA34620.1; ALT\_SEQ.  
 DR PIR; A56577; A56577.  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 8.  
 KW Microtubule; Repeat; Phosphorylation.  
 FT CHAIN ? 2459 MAP1 LIGHT CHAIN LC1.

FT	REPEAT	1869	1885	MAP1B 1.
FT	REPEAT	1886	1902	MAP1B 2.
FT	REPEAT	1903	1919	MAP1B 3.
FT	REPEAT	1920	1936	MAP1B 4.
FT	REPEAT	1937	1953	MAP1B 5.
FT	REPEAT	1954	1970	MAP1B 6.
FT	REPEAT	1988	2004	MAP1B 7.
FT	REPEAT	2005	2021	MAP1B 8.
FT	REPEAT	2022	2038	MAP1B 9.
FT	REPEAT	2039	2055	MAP1B 10.
FT	DOMAIN	559	1035	GLU-RICH.
FT	DOMAIN	588	786	LYS-RICH (HIGHLY BASIC, CONTAINS MANY KKEE AND KKEI/V REPEATS).
FT	DOMAIN	2224	2312	LYS-RICH.
FT	CONFLICT	127	127	M -> V (IN REF. 1).
FT	CONFLICT	140	140	T -> S (IN REF. 1).
FT	CONFLICT	2112	2112	R -> K (IN REF. 3).
FT	CONFLICT	2169	2169	L -> I (IN REF. 3).
SQ	SEQUENCE	2459	AA; 269497	MW; 2E3F6872DEDB8BA2 CRC64;

Query Match 5.0%; Score 291.5; DB 1; Length 2459;  
 Best Local Similarity 20.0%; Pred. No. 0.0013;  
 Matches 220; Conservative 167; Mismatches 456; Indels 255; Gaps 44;

Qy	30	TEPEDEEDEEEEEDEEDEDLEELEVLERKPAAGLSAAAVPPAAAAPLLDFFSSDSVPPA	89
		:      : :       :  : :          : :   :	
Db	1008	SEEEGEEEEEDKAEDAREEDHEPDKTE-----AEDYVMAVVDKAAEAGVTEDQYDFL---	1058
Qy	90	PRGPLPAAPPAAPERQPSWERSPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGAS	149
		:    :     ::	
Db	1059	-----GTPAKQ-----PGVQSPSREPASSIHDETLPGGSESEAT-----AS	1094
Qy	150	PLAEPAAPSTPAAPKRRGSGSVDETFLFALPAASEFPVIPSSAEKIMDLMEQPGN---TVS	206
		: :     :	
Db	1095	DEENREDQPEEFTAT----SGYTQST---IEISSEPTPMDDEMSTPRDVMTDETNNETES	1147
Qy	207	SGQE-----DFPSVLLETAASLP---SLSPLS-----TVSFKEHGYLGNL SAVSSSEGTIE	254
		:         :      :   :   :	
Db	1148	PSQEFVNITKYESSLYSQEYSKPVVASFNLSDGSKTDATDGRDYNASASTISPPSSMEE	1207
Qy	255	ETLNEAS-----KELPERATNPFVNRDLAEF--SELEYSEMG---	289
		: :::   : :   :   :   :	
Db	1208	DKFSKSALRDAYRPEETDVKTGAELDIKDVSDERLSPAKSPSLSPSPSPPIEKTPLGERS	1267
Qy	290	-----SSFKGSPKGESAILVENTKEEVIVR----SKDKEDL-----VCSAALHSP	330
		:     :  : :  : :    :::     :    :	
Db	1268	VNFSLTPNEIKASAEGEATAVVSPGVTQAVVEEHCASPEEKTLEVVSPSQSVTGSAGHTP	1327
Qy	331	-QESPVGKEDRVVSPEKTMDFNEMQMSVVAPVREEYADFK-----PFEQAWEVKD	380
		:   :: :     : :      : :     : :	
Db	1328	YYQSPTDEKSSHLPT EVT-----ENQAQVPVSFEFTEAKDENERSISSIPMDE--PVPD	1378
Qy	381	TYEGSRDVLAAARANVESKVDRKCLED--SLEQKSLGKDS----EGRNEDASFPSTPEPVK	434
		:   : :      :   :     :	
Db	1379	SESPIEKVLSPLRSPPLIGSESAYEDFLSADDKALGRRSESPFEGKNGKQGFSDKESPV	1438
Qy	435	DSSRAYITCASFTSATESTTANTFPLLEDHTSENKT-----DEKKI-----	475

Db	1439	D-----LTSDLYQDKQEEKRAGFIPIKEDFSPEKKASDAEIMSSQSALALDERKLGGDGS	1493
Qy	476	-----EERKAQIITEKTSNPKTSNPFVLAVQDSEADYVTTDTLSKVTEAAV--	520
Db	1494	PTQVDVSQFGSFKEDTKMSISEGTVSDKSATPVDEGAEDT---YSHMEGVASVSTASVAT	1550
Qy	521	SNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAIQES-LYPTAQLCP---S	576
Db	1551	SSFPEPTTDD-VSPSLHAEVGSPPHSTEVDDSLSVSVVQTPTTFQETEMSPSKEECPRPMS	1609
Qy	577	FEEAEATP-----SPVLPDIVMEAPLNSLL-----PSAGASVVQ	610
Db	1610	ISPPDFSPKTAKSRTVPVQDHRSEQSSMSIEFGQESPEHSLAMDFSRQSPDHPTVGAGMLH	1669
Qy	611	PSVSPLEAPPPVSYD-----SIKLEPENPPPYEEAMNVALKALGTKEGIKEPESFN	661
Db	1670	ITEN---GPTEVDYSPSDIQDSSLHKIPPTTEPSYTDNDLS-ELISVSQVEASPSTSS	1725
Qy	662	AAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDSSPESEP	721
Db	1726	AHTPS-----QIASPLQEDTLSDVPPRDMSLYASLASEKVQSLEGEKL----SPKSDI	1775
Qy	722	VDLFSDDSIPEVPQTQEEAVMLMKESLTVSETVAQHKEERLSASPQELGKPYL-ESFQP	780
Db	1776	SPLTPRESSPTYSPGFSDSTSGAKES-TAAYQTSSSPPIDAAAAEPYGRSSMLFDTMQH	1834
Qy	781	NLHSTKDAASNDIPTLTKEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIE	840
Db	1835	HLALSRLTTSSV----EKDNGGKTPGDFNYAYQKPESTTESPDEEDYDYE--SHEKTIQ	1888
Qy	841	IIDEFPTFVSAKDDSPK----LAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFKN	896
Db	1889	AHDVGGYYEKTERTIKSPCDSGYSYETIEKTKTP----EDGGYS-----CEITEKT	1937
Qy	897	IYPKDEVHVSDEFSENRSSVSKAS-----ISPSNVSALEPQTEMGSIVKSKSL	944
Db	1938	TRTPEEGGYSYEISEKTRTRTEPVSGYTYEKTERRRLDDISNGYDDTEDGGHTLGDCSY	1997
Qy	945	TKEAEKKLPDTEKEDRS	962
Db	1998	SYETTEKITSFPESESYS	2015

RESULT 14

MAPB\_HUMAN

ID MAPB\_HUMAN STANDARD; PRT; 2468 AA.  
AC P46821;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain  
DE LC1].  
GN MAP1B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=95104835; PubMed=7806212;  
 RA Lien L.L., Feener C., Fischbach N., Kunkel L.M.;  
 RT "Cloning of human microtubule-associated protein 1B and the  
 RT identification of a related gene on chromosome 15.";  
 RL Genomics 22:273-280(1994).  
 CC -!- FUNCTION: The function of brain MAPS is essentially unknown.  
 CC     Phosphorylated MAP1B may play a role in the cytoskeletal changes  
 CC     that accompany neurite extension. Possibly MAP1B Binds to at least  
 CC     two tubulin subunits in the polymer, and this bridging of subunits  
 CC     might be involved in nucleating microtubule polymerization and in  
 CC     stabilizing microtubules.  
 CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate  
 CC     with MAP1A and MAP1B proteins.  
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC     KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC     responsible for the binding of MAP1B to microtubules.  
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated  
 CC     from MAP1B by proteolytic processing. It is free to associate with  
 CC     both MAP1A and MAP1B. It interacts with the amino-terminal region  
 CC     of MAP1B (By similarity).  
 CC -!- SIMILARITY: TO MAP1A.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L06237; AAA18904.1; -.  
 DR Genew; HGNC:6836; MAP1B.  
 DR MIM; 157129; -.  
 DR GO; GO:0005875; C:microtubule associated complex; TAS.  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 6.  
 KW Microtubule; Repeat; Phosphorylation.  
 FT CHAIN           ?   2468           MAP1 LIGHT CHAIN LC1.  
 FT REPEAT       1878   1894           MAP1B 1.  
 FT REPEAT       1895   1911           MAP1B 2.  
 FT REPEAT       1912   1928           MAP1B 3.  
 FT REPEAT       1929   1945           MAP1B 4.  
 FT REPEAT       1946   1962           MAP1B 5.  
 FT REPEAT       1963   1979           MAP1B 6.  
 FT REPEAT       1997   2013           MAP1B 7.  
 FT REPEAT       2014   2030           MAP1B 8.  
 FT REPEAT       2031   2047           MAP1B 9.  
 FT REPEAT       2048   2064           MAP1B 10.  
 FT DOMAIN       589     790           LYS-RICH (HIGHLY BASIC, CONTAINS MANY  
 FT                                   KKEE AND KKEI/V REPEATS).  
 SQ SEQUENCE   2468 AA; 270618 MW; 540839CBDF09D461 CRC64;

Query Match 4.9%; Score 288; DB 1; Length 2468;  
Best Local Similarity 20.2%; Pred. No. 0.0017;  
Matches 251; Conservative 166; Mismatches 441; Indels 386; Gaps 56;

```
Qy      13 STDSPPR--PPPAFKYQFVTEPEDEEDEEEE-----EED-----EEEDDED 50
      :|| |: | : :|||:::|: | :|| | |
Db      625 ATDVKPKAAKEKTVKKETKVKPEDKKEEKEKPKKEVAKKEDKTPIKKEEKPKKEEVKKEV 684

Qy      51 LEELEVLERKPAAGLSAAAVPPAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWER 110
      :|:: |:| | : :| : : : :|: | :|: |
Db      685 KKEIKKEEKK-----EPKKEVKKETPPKEVKKEVKKEEKKEVKKE---EK 726

Qy     111 SPAAPAPSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAPPSTPAAPKRRG-- 168
      | | | : | | : | | | | : | | | :|
Db      727 EPKKEIKKLPKDAKKSSTPLSEAKKPAALKPKVPKKEESVKKDSVA----AGKPKEKGKI 782

Qy     169 -----SGSVDETLFAL-----PA-----ASEVIPSSAEKIMDLME 199
      | | : | | | : | : | : | |
Db      783 KVIKKEGKAAEAVAAVGTGATTAAVMAAAGIAAIGPAKELEAERSLMSSPEDLTkdFEE 842

Qy     200 QPGNTVSSGQEDFPSV-LLETAASLPSLSPLSTVSFKEHGYLGNLSAVSSSEGTIEETLN 258
      | : : | : | : | | : : : | | | |
Db      843 LKAEVDVTKDIKPQLELIEDEEKLKETEPVEAYVIQKEREVTKGPAESPDEG-ITTTEG 901

Qy     259 EASKELPERATNPFVNRDLAEFSELEYSEMSSFKGSPKGESAILVENTK-EEVIVRSKD 317
      | | | : : : : | : | : | | : | : || :|
Db      902 EGECEQTPEELEPVEKQGVDDIEKFE--DEGAGFEES--SETGDYEEKAETEEAEEPEED 957

Qy     318 KEDLVC-SAALHSP-----QESPVGKEDRVVSPEKTMDIFNEMQMSV 358
      |: || ||: ||| :|| :|| | : || |
Db      958 GEEHVCVSASKHSPTEDDEESAKAEADAYIREKRESVASGDDRA---EEDMD---EAIEKG 1011

Qy     359 VAPVREEYADFKPFQAWEVK-DTYEGSRDVLAAARANVESKVDRKCLEDLSLEQ----- 410
      | || || : :|| : : : || : : | | : || :|
Db     1012 EAEQSEEEADEE--DKAEDAREEEYEPEK--MEAEDYVMAVVDKAAEAGGAEEQYGLT 1067

Qy     411 --KSLGKDSEGR-----NEDASFPSTPEPVKDSSRAYITCASFTSATESTT 454
      | || | || :| : | | :| : : :| :| :
Db     1068 PTKQLGAQSPGREPASSIHDETLPGGSESEATASDEENREDQPEEFTATSGYTQSTIEIS 1127

Qy     455 ANTFPL-----LEDHTSENKTDEK-----KIEERKAQIITEKTS-PKTSNPFLVAV 499
      : |: : | |: :| : | : : : :| | |
Db     1128 SEPTPMDEMSTPRDVMSDETNNREETESPSQEFVNITKYESSLYSQEYSKPADVTPLNGFS 1187

Qy     500 QDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEACESELNEATGTKIAYETKVDL--- 556
      : |: | || | : : | : | : | | :| :| :|:
Db     1188 EGSKTD--ATDGKDYNASASTISPPSSMEED--KFSRSALRDAYCEVKASTTLDIKDS 1242

Qy     557 --VQTSEAIQESLYPTAQLCP-----SFE-----EAEATSPVLPDIV 592
      :|| : | |: | :| :|| :|| |::
Db     1243 ISAVSSEKVSPSKSPSLSPSPSPLEKTPLGERSVNFSLTPNEIKVSAEAEVAPVSPEVT 1302

Qy     593 ME-----APLNSLLPSAG-----ASVVQ----- 610
      | :| :| || | :|
Db     1303 QEVVEEHCASPEDKTLEVVSQS SVTGSAGHTPYYSPTDEKSSHLPTVIEKPPAVPVS 1362

Qy     611 -----PSVSPLEAPPVSYDSIK--LEPENPPP-----YEEAMNVALKALG 649
```

```

Db      1363 FEFSDAKDENERASVSPMDEPVPDSESPIEKVLSPLRSPPLIGSESAYESFLSADDKASG 1422
QY      650 TKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPD-FSNYSEIA-----KF 701
      : |:: | | : | | | | : |
Db      1423 -----RGAESPF-----EEKSGKQGSPDQVSPVSEMTSTSLYQDKQ 1458
QY      702 EKSVPHEAELVEDSSPE-----SEPVDLFSDDSIPEVPQTQ-----EEAV 741
      | : | : | | | | | : | : | : | | | : |
Db      1459 EGKSTDFAPIKEDFGQEKKTDVVEAMSSQPALALDERKLGDVSPTQIDVSQFGSFKEDTK 1518
QY      742 MLMKE-----SLTEVSETVAQ----HKEERLSASPQELG-----KPYLESFQPNLHSTK 786
      | : | | | | | | | : | : | : | : | : |
Db      1519 MSISEGTVSDKSATPVDEGVAEDTYSHMEGVASVSTASVATSSFPEPTTDDVSPSLH--- 1575
QY      787 DAASNDIPTLTK-KEKISLQMEEFNTAIYSNDDLSSKEDKIKESETFSDSSPIEIIDEF 845
      | | | : : : : | : : : | | :
Db      1576 --AEVGSPHSTEVDDSLSVSVVQTPTT-FQETEMSPSKEECPR----- 1615
QY      846 PTFVSAKDDSPKLAKEYTDLE--VSDKSEIANIQSGADSLPCLELPCDLSFKN----- 896
      | : | | | | | : : | : : : : | : | | : :
Db      1616 PMSISPPDFSPKTAKSRTFPVQDHRSEQSSM-SIEFGQES-PEQSLAMDFSRQSPDHPTVG 1673
QY      897 -----IYPKDE-VHVSDEFSENRRSSVSKASISPSNV 926
      | | : | : | | : | |
Db      1674 AGVLHITENGPTVDYSPSDMQDSSLSHKIPPEEPSYTDNDLSELISVSQVEASPSTS 1733
QY      927 SALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKEDRSLSAVLSAE 970
      || | : : | : : : | : | | | : : |
Db      1734 SAHTP-SQIASPLQEDTLSDVAPPR-----DMSLYASLTSE 1768

```

# RESULT 15

## MAPB\_MOUSE

```

ID  MAPB_MOUSE      STANDARD;      PRT;  2464 AA.
AC  P14873;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-APR-1990 (Rel. 14, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Microtubule-associated protein 1B (MAP 1B) (MAP1.2) (MAP1(X))
DE  [Contains: MAP1 light chain LC1].
GN  MAP1B OR MTAP1B OR MTAP5.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A., AND DOMAIN.
RC  STRAIN=Swiss Webster; TISSUE=Brain;
RX  MEDLINE=90094539; PubMed=2480963;
RA  Noble M., Lewis S.A., Cowan N.J.;
RT  "The microtubule binding domain of microtubule-associated protein
RT  MAP1B contains a repeated sequence motif unrelated to that of MAP2
RT  and tau.";
RL  J. Cell Biol. 109:3367-3376(1989).
CC  -!- FUNCTION: The function of brain MAPS is essentially unknown.
CC  Phosphorylated MAP1B may play a role in the cytoskeletal changes

```

CC that accompany neurite extension. Possibly MAP1B Binds to at least  
 CC two tubulin subunits in the polymer, and this bridging of subunits  
 CC might be involved in nucleating microtubule polymerization and in  
 CC stabilizing microtubules.

CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate  
 CC with MAP1A and MAP1B proteins.

CC -!- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC responsible for the binding of MAP1B to microtubules.

CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated  
 CC from MAP1B by proteolytic processing. It is free to associate with  
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region  
 CC of MAP1B.

CC -!- SIMILARITY: TO MAP1A.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X51396; CAA35761.1; -.  
 DR PIR; S07549; QRMSPL.  
 DR MGD; MGI:1306778; Mtap1b.  
 DR GO; GO:0016358; P:dendrite morphogenesis; IMP.  
 DR GO; GO:0001578; P:microtubule bundling; IMP.  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 7.  
 KW Microtubule; Repeat; Phosphorylation.

FT	CHAIN	?	2464	MAP1 LIGHT CHAIN LC1.
FT	REPEAT	1874	1890	MAP1B 1.
FT	REPEAT	1891	1907	MAP1B 2.
FT	REPEAT	1908	1924	MAP1B 3.
FT	REPEAT	1925	1941	MAP1B 4.
FT	REPEAT	1942	1958	MAP1B 5.
FT	REPEAT	1959	1975	MAP1B 6.
FT	REPEAT	1993	2009	MAP1B 7.
FT	REPEAT	2010	2026	MAP1B 8.
FT	REPEAT	2027	2043	MAP1B 9.
FT	REPEAT	2044	2060	MAP1B 10.
FT	DOMAIN	589	787	LYS-RICH (HIGHLY BASIC, CONTAINS MANY
FT				KKEE AND KKEI/V REPEATS).

SQ SEQUENCE 2464 AA; 270408 MW; FBD3DD99CFDBDA87 CRC64;

Query Match 4.9%; Score 284; DB 1; Length 2464;  
 Best Local Similarity 20.7%; Pred. No. 0.0025;  
 Matches 233; Conservative 159; Mismatches 424; Indels 312; Gaps 51;

QY 31 EPEDEEDEEEEEDEEED-----DEDLEELE-----VLERKPAAG-----LSAAAVP 71  
 | | :| |||: || : | | | :| :| :|  
 Db 1009 EAEQSEEEGEEEDKAEDAREEGYEPDKTEAEDYVMAVADKAAEAGVTEEQYGYLGTS AKQ 1068

QY 72 PAAAPLLDFSS----DSVPPAPRGPLPAAPPAAPERQP-----SWERSPAAPA 116  
 | :| : :| ::| | : | | :|



Db	1069	PGIQSPSPREPASSIHDETLPGGSESEATASDEENREDQPEEFTATSGYTQSTIEISSEPT	1128
Qy	117	P---SLPPAAAVLPSKLPEDDEPPARP-----PPPPPAGASPLAEPA--	155
Db	1129	PMDEMSTPRDVMSDETNNNEETESPSQEFVNITKYESSLYSQEYSKPAVASFNGLSEGSKT	1188
Qy	156	-----APPSTPAAPKRRGSGSVDETLFALPA-----ASEPVIPSSAEKIMDLMEQ	200
Db	1189	DATDGKDYNASASTISPP-----SSMEEDKFSKSALRDAYCSEEKELKASAE--LDIKDV	1241
Qy	201	PGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFKEHGYLGNLSAVSSSEGTTIEETLNEA	260
Db	1242	SDERLS-----PAKSPSLSP-----SPPSPIEKT----	1265
Qy	261	SKELPERATNPVFNRLAEFS----ELEYSEMGSFFKGSPKGESAILVEN--TKEEVIVR	314
Db	1266	--PLGERSVN-----FSLTPNEIKVSAEGEARSVSPGVQTQAVVEEHCASPEEKTLE	1314
Qy	315	SKDKEDLVCSAALHSP--QESPVGKEDRVVSPEKTMDFNEMQMSVVA--PVREEYADFK--	370
Db	1315	VVSPSQSVTGSAGHTPYQSP-----TDEKSSHLPTVEVSENAQAVPVSFEFSEAKDE	1366
Qy	371	-----PFEQAWEVKDTYEGSRDVLAAARANVESKVDRKLED--SLEQKSLGKDS----	417
Db	1367	NERASLSPMDE--PVPDSESPVEKVLSPLRSPPLLGSESPYEDFLSADSKVLGRRSESPF	1424
Qy	418	EGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPILLEDHTSENKT-----	470
Db	1425	EGKNGKQGFPDRESPVSDLT----STGLYQDKQEEKSTGFIPIKEDFGPEKKTSDVETMS	1480
Qy	471	-----DEKKI-----EERKAQIITEKTSPKTSNPFVAVQDSEAD	505
Db	1481	SQSALALDERKLGGDVSPTQIDVSQFGSFKEDTKMSISEGTVSDKSATPVDEGV--AEDT	1538
Qy	506	YVTTDTLSKVTEAAV--SNMPEGLTPDLVQEACESELNEATGTKIAYETKVDLVQTSEAI	563
Db	1539	YSHMEGVASVSTASVATSSFPEPTTDD-VSPSLHAEVGSHPSTEVDDSLSVSVVQTPTTF	1597
Qy	564	QES-LYPTAQLCP---SFEEAEATP-----SPVLPDIVMEAPLNSLL--	601
Db	1598	QETEMSPSKEECPRPMSSISPPDFSPKTAKSRTPVQDHRSEQSSMSIEFGQESPEHSFAMD	1657
Qy	602	-----PSAGASVVQPSVSPLEAPPPVSYD-----SIKLEPENPPPYEEAMNVA	644
Db	1658	FSRQSPDHPTLGASVLHITEN---GPTEVDYSPCDIQDSSSLSHKIPPTTEEPSYTDNDLS	1714
Qy	645	LKALGTKEGIKEPESFNAAVQETEAPYISACDLIKETKLSTEPSPDFSNYSEIAKFEKS	704
Db	1715	-ELISVSQVEASPSTSSAHTPS-----QIASPLQEDTLSDVPPPREMSLYASLA-----	1762
Qy	705	VPEHAELVEDSSPESEPVDLFSDDSIPEVPQTQEEAVML---MKESLTVSETVAQHKE	760
Db	1763	----SEKVQ--SLEGEKLSPKSDIS----PLTPRESSPLYSPGFSDDSTSAAKETAAAH--	1810
Qy	761	ERLSASP---QELGKPY-----LESFQPNLHSTKDAASNDIPTLTKEKISLQMEEFN	810
Db	1811	-QASSPPIIDAATAEPYGFRRSSMLFDTMQHHLALNRDLTSSV---EKDSGGKTPGDFN	1865

Qy 811 TAIYSNDDLSSKEDKIKESETFSDSSPIEIIDFPTFVSAKDDSPK----LAKEYTDLE 866  
 | :: | ::: : | | | : : | : |  
 Db 1866 YAYQKPENAAGSPDEEDYDYE--SQEKTIRTHDVVRYYYEKTERTIKSPCDSGYSYETIE 1923

Qy 867 VSDKSEIANIQSGADSLPCLELPCDLSFKNIYPKDEVHVSDEFSENRSSVSKAS----- 920  
 : |: : | |:: | :| | | | : : |  
 Db 1924 KTTKTP----EDGG-----YTCEITEKTTTRTPEEGGYSEISEKTTRTPEVSGYTYEK 1972

Qy 921 -----ISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSDTEKEDRS 962  
 : : : : | : | : | :| | | |  
 Db 1973 TERSRRLDDISNGYDDTEDGGHTLGDCSYSYETTEKITSFPESESYS 2020

Search completed: September 29, 2004, 18:14:37  
 Job time : 27.6724 secs